

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 21:16:45 ; Search time 597.656 Seconds  
(without alignments)  
2026.885 Million cell updates/sec

Title: US-10-808-187A-2471  
Perfect score: 25  
Sequence: 1 cgaacgctgtagcttcaaaatct 25

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_cm.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	626	14	AY451928 SARS CORO
2	25	100.0	626	14	AY451929 SARS CORO
3	25	100.0	626	14	AY451930 SARS CORO
4	25	100.0	626	14	AY451931 SARS CORO
5	25	100.0	626	14	AY451932 SARS CORO
6	25	100.0	626	14	AY451933 SARS CORO
7	25	100.0	626	14	AY451934 SARS CORO
8	25	100.0	626	14	AY451935 SARS CORO
9	25	100.0	626	14	AY451936 SARS CORO
10	25	100.0	626	14	AY451937 SARS CORO
11	25	100.0	626	14	AY451938 SARS CORO
12	25	100.0	626	14	AY451939 SARS CORO
13	25	100.0	626	14	AY451940 SARS CORO
14	25	100.0	626	14	AY451941 SARS CORO
15	25	100.0	626	14	AY451942 SARS CORO
16	25	100.0	626	14	AY451943 SARS CORO
17	25	100.0	626	14	AY451944 SARS CORO
18	25	100.0	626	14	AY451945 SARS CORO
19	25	100.0	677	14	AY443086S05

20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45

AY534764 SARS CORO  
AY534759 SARS CORO  
AY304492 SARS CORO  
AY286320 SARS CORO  
AY463060 SARS CORO  
AY394999 SARS CORO  
AY395000 SARS CORO  
AY395001 SARS CORO  
AY395002 SARS CORO  
AY394977 SARS CORO  
AY394985 SARS CORO  
AY338174 SARS CORO  
AY338175 SARS CORO  
AY348314 SARS CORO  
AY559094 SARS CORO  
AY463059 SARS CORO  
AY395004 SARS CORO  
AY394978 SARS CORO  
AY394979 SARS CORO  
AY394982 SARS CORO  
AY559086 SARS CORO  
AY394988 SARS CORO  
AY559082 SARS CORO  
AY394983 SARS CORO  
AY283795 SARS CORO  
AY394980 SARS CORO

ALIGNMENTS

RESULT 1  
AY451928  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AY451928  
SARS coronavirus TW-HPI isolate TW-HPI\_SCI8 replicase 1B gene,  
partial cds.  
AY451928  
AY451928.1 GI:42741328  
SARS coronavirus TW-HPI  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE  
AUTHORS

Yang, J.-Y., Chen, H.-Y., and Chen, Y.-M. Arthur.

Molecular Epidemiology of SARS in Taiwan - Development of a Simple

Method for Tracing the Origin and Dissemination of SARS

Unpublished

2 (bases 1 to 626)

Yang, J.-Y., Chen, H.-Y., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,

Direct Submission

Submitted (28-OCT-2003) National Yang-Ming University, AIDS

Prevention and Research Center, 155, Li-Nong St., Sec.2, Beitou,

Taipei, Taiwan 112, R.O.C.

Location/Qualifiers

1. .626

/organism="SARS coronavirus TW-HPI"

/mol\_type="genomic RNA"

/isolate="TW-HPI\_SCI8"

/db\_xref="taxon:264371"

<1..>626

/codon\_start=1

/product="replicase 1B"

/protein\_id="AA544814.1"

/db\_xref="GI:42741329"

/translation="ISMATNYDLISVVNARLAKHYVYIGDPAQLPAPRTLTKGTLEP

EYFNSVCLMKTTIGDPMFLGTCRCPAETVDVTSALVYNKLUKHKDKSAQCQFMFYK

GVTHDVSAINRPOIVGVRELTENPAWKAVFISPNYSONAVASKILGLPTQTVDS

SGSGEYDVYVFTQTETAHSCNVNRFNVAITRAKIGILCIMSDDLID"

ORIGIN

```
Query Match      100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451

RESULT 2
AY451929
LOCUS SARS coronavirus TW-HP2 RNA linear VRL 28-FEB-2004
DEFINITION partial cds.
ACCESSION AY451929
VERSION AY451929.1 GI:42741330
KEYWORDS
SOURCE
ORGANISM SARS coronavirus TW-HP2
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
JOURNAL Method for Tracing the Origin and Dissemination of SARS
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
Direct Submission
Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1. .626
/organism="SARS coronavirus TW-HP2"
/mol_type="genomic RNA"
/isolate="TW-HP2_SC18"
/db_xref="taxon:264372"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44815.1"
/db_xref="GI:42741330"
/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLTKGTLEP
EYFNSVCLMKRTIGDPMFLGTCRCPPAEIVDTVSALYVDNKLKAHKDKSAQCCKMFYK
GVITHDVSSAINRPOIGVVRFLTRNPAWRKAVFISPNYSONAVASKILGLPTQTIVDS
SQGSYDYVITQTTFETAHSCNVNRFNVAITRAKIGILCIMSDRDLVD"

ORIGIN
Query Match      100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451

RESULT 4
AY451931
LOCUS SARS coronavirus TW-HP4 isolate TW-HP4_SC18 linear VRL 28-FEB-2004
DEFINITION partial cds.
ACCESSION AY451931
VERSION AY451931.1 GI:42741334
KEYWORDS SARS coronavirus TW-HP4
SOURCE SARS coronavirus TW-HP4
ORGANISM SARS coronavirus TW-HP4
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
JOURNAL Method for Tracing the Origin and Dissemination of SARS
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
Direct Submission
Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1. .626
/organism="SARS coronavirus TW-HP4"
/mol_type="genomic RNA"
/isolate="TW-HP4_SC18"
/db_xref="taxon:264374"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44817.1"

CDS
```

```
/db_xref="GI:42741335"
/translation="ISMATNYDLVVNARLRAKHVYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCRMKTIGPMFLGTCRCRPAEIVDTVSALVYDNKLAHKDKSAQCFKMFYK
GWITHDVSAINRPOIGVREFLTRNPARKAVFISPYNSQNAVASKILGLPTQTVD
SQGSEYDVYVIFTQTETAHSCNVNRFNVAITRAKIGILCIMSRLDLYD"
ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451
RESULT 5
AY451932 SARS coronavirus TW-JC2 isolate TW-JC2_SCI18 replicase 1B gene,
LOCUS partial cds.
DEFINITION
ACCESSION AY451932
VERSION AY451932.1 GI:42741336
KEYWORDS
SOURCE SARS coronavirus TW-JC2
ORGANISM SARS coronavirus TW-JC2
REFERENCE
1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1..626
/organism="SARS coronavirus TW-JC2"
/mol_type="genomic RNA"
/isolate="TW-JC2_SCI18"
/db_xref="taxon:264375"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44818.1"
/db_xref="GI:42741337"
/translation="ISMATNYDLVVNARLRAKHVYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCRMKTIGPMFLGTCRCRPAEIVDTVSALVYDNKLAHKDKSAQCFKMFYK
GWITHDVSAINRPOIGVREFLTRNPARKAVFISPYNSQNAVASKILGLPTQTVD
SQGSEYDVYVIFTQTETAHSCNVNRFNVAITRAKIGILCIMSRLDLYD"
CDS
1..626
/organism="SARS coronavirus TW-JC2"
/mol_type="genomic RNA"
/isolate="TW-JC2_SCI18"
/db_xref="taxon:264375"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44818.1"
/db_xref="GI:42741337"
/translation="ISMATNYDLVVNARLRAKHVYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCRMKTIGPMFLGTCRCRPAEIVDTVSALVYDNKLAHKDKSAQCFKMFYK
GWITHDVSAINRPOIGVREFLTRNPARKAVFISPYNSQNAVASKILGLPTQTVD
SQGSEYDVYVIFTQTETAHSCNVNRFNVAITRAKIGILCIMSRLDLYD"
ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451
RESULT 6
AY451933 SARS coronavirus TW-KC1 isolate TW-KC1_SCI18 replicase 1B gene,
LOCUS partial cds.
DEFINITION
ACCESSION AY451933
```

```
AY451933.1 GI:42741338
SARS coronavirus TW-KC1
SARS coronavirus TW-KC1
Viruses; sRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
1 (bases 1 to 626)
Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1..626
/organism="SARS coronavirus TW-KC1"
/mol_type="genomic RNA"
/isolate="TW-KC1_SCI18"
/db_xref="taxon:264376"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44819.1"
/db_xref="GI:42741339"
/translation="ISMATNYDLVVNARLRAKHVYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCRMKTIGPMFLGTCRCRPAEIVDTVSALVYDNKLAHKDKSAQCFKMFYK
GWITHDVSAINRPOIGVREFLTRNPARKAVFISPYNSQNAVASKILGLPTQTVD
SQGSEYDVYVIFTQTETAHSCNVNRFNVAITRAKIGILCIMSRLDLYD"
ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451
RESULT 7
AY451934 SARS coronavirus TW-KC3 isolate TW-KC3_SCI18 replicase 1B gene,
LOCUS partial cds.
DEFINITION
ACCESSION AY451934
VERSION AY451934.1 GI:42741340
KEYWORDS SARS coronavirus TW-KC3
SOURCE SARS coronavirus TW-KC3
ORGANISM Viruses; sRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
1 (bases 1 to 626)
Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1..626
/organism="SARS coronavirus TW-KC3"
```

```

/mol_type="genomic RNA"
/isolate="TW-KC3_SC18"
/db_xref="taxon:264377"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44820.1"
/db_xref="GI:42741341"
/translation="ISMATNYDLSVNNARLRAKHVYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCLMTIGDMPFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCFFRMFYK
GVITHDVSAINRPOIGVREFLTRNPAWRKAVFISPNYQNAVASKILGLPTQTVD
SQSGEYDYVIFTQTETAHSCNVNRFVVAITRAKIGILCINMSDRDLYD"

ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451

```

```

RESULT 8
AY451935
LOCUS SARS coronavirus TW-GD1 isolate TW-GD1_SC18 replicase 1B gene,
partial cds.
VERSION AY451935.1 GI:42741342
KEYWORDS SARS coronavirus TW-GD1
SOURCE SARS coronavirus TW-GD1
ORGANISM SARS coronavirus TW-GD1
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
Location/Qualifiers
1..626
/organism="SARS coronavirus TW-GD1"
/mol_type="genomic RNA"
/isolate="TW-GD1_SC18"
/db_xref="taxon:264378"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44821.1"
/db_xref="GI:42741343"
/translation="ISMATNYDLSVNNARLRAKHVYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCLMTIGDMPFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCFFRMFYK
GVITHDVSAINRPOIGVREFLTRNPAWRKAVFISPNYQNAVASKILGLPTQTVD
SQSGEYDYVIFTQTETAHSCNVNRFVVAITRAKIGILCINMSDRDLYD"

CDS
Query Match 100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451

```

```

RESULT 10
AY451937
LOCUS SARS coronavirus TW-GD3 isolate TW-GD3_SC18 replicase 1B gene,
partial cds.
VERSION AY451937
KEYWORDS SARS coronavirus TW-GD3
SOURCE SARS coronavirus TW-GD3
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.

```

```

RESULT 9
AY451936
LOCUS SARS coronavirus TW-GD2 isolate TW-GD2_SC18 replicase 1B gene,
partial cds.
VERSION AY451936
KEYWORDS SARS coronavirus TW-GD2
SOURCE SARS coronavirus TW-GD2
ORGANISM SARS coronavirus TW-GD2
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
Location/Qualifiers
1..626
/organism="SARS coronavirus TW-GD2"
/mol_type="genomic RNA"
/isolate="TW-GD2_SC18"
/db_xref="taxon:264379"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44822.1"
/db_xref="GI:42741345"
/translation="ISMATNYDLSVNNARLRAKHVYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCLMTIGDMPFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCFFRMFYK
GVITHDVSAINRPOIGVREFLTRNPAWRKAVFISPNYQNAVASKILGLPTQTVD
SQSGEYDYVIFTQTETAHSCNVNRFVVAITRAKIGILCINMSDRDLYD"

ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451

```

```

RESULT 10
AY451937
LOCUS SARS coronavirus TW-GD3 isolate TW-GD3_SC18 replicase 1B gene,
partial cds.
VERSION AY451937
KEYWORDS SARS coronavirus TW-GD3
SOURCE SARS coronavirus TW-GD3
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.

```



```

TITLE      Direct Submission
JOURNAL    Submitted (28-OCT-2003) National Yang-Ming University, AIDS
           Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
           Taipei, Taiwan 112, R.O.C.
FEATURES   source
           1. .626
           /organism="SARS coronavirus TW-GD3"
           /mol_type="genomic RNA"
           /isolate="TW-GD3_SCI8"
           /db_xref="taxon:264380"
           <1. .>626
           /codon_start=1
           /product="replicase 1B"
           /protein_id="AA544823.1"
           /db_xref="GI:42741347"
           /translations="ISMATNYDLSVNNARLAKHYVYIGDPAQLPAPRTLLTKGTLEP
           EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCQKMFYK
           GVITHDVSSAINRPOIGVREFLTRNPARKAVFISPYNSONAVASKILGLPTQTVD
           SQSEYDYVIFTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN
Query Match      100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db      427 CAGAACGCTGTAGCTTCAAAAATCT 451

RESULT 11
AY451938
LOCUS    SARS coronavirus TW-GD4 isolate TW-GD4_SCI8 replicase 1B gene,
partial cds.
ACCESSION AY451938.1 GI:42741348
VERSION   SARS coronavirus TW-GD4
KEYWORDS  Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
ORGANISM  1 (bases 1 to 626)
           Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Molecular Epidemiology of SARS in Taiwan - Development of a Simple
           Method for Tracing the Origin and Dissemination of SARS
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS   Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Direct Submission
TITLE     Submitted (28-OCT-2003) National Yang-Ming University, AIDS
           Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
           Taipei, Taiwan 112, R.O.C.
JOURNAL   Location/Qualifiers
FEATURES   source
           1. .626
           /organism="SARS coronavirus TW-GD4"
           /mol_type="genomic RNA"
           /isolate="TW-GD4_SCI8"
           /db_xref="taxon:264381"
           <1. .>626
           /codon_start=1
           /product="replicase 1B"
           /protein_id="AA544824.1"
           /db_xref="GI:42741349"
           /translations="ISMATNYDLSVNNARLAKHYVYIGDPAQLPAPRTLLTKGTLEP
           EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCQKMFYK
           GVITHDVSSAINRPOIGVREFLTRNPARKAVFISPYNSONAVASKILGLPTQTVD
           SQSEYDYVIFTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN
Query Match      100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db      427 CAGAACGCTGTAGCTTCAAAAATCT 451

RESULT 12
AY451939
LOCUS    SARS coronavirus TW-GD5 isolate TW-GD5_SCI8 replicase 1B gene,
partial cds.
ACCESSION AY451939.1 GI:42741350
VERSION   SARS coronavirus TW-GD5
KEYWORDS  Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
ORGANISM  1 (bases 1 to 626)
           Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Molecular Epidemiology of SARS in Taiwan - Development of a Simple
           Method for Tracing the Origin and Dissemination of SARS
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS   Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Direct Submission
TITLE     Submitted (28-OCT-2003) National Yang-Ming University, AIDS
           Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
           Taipei, Taiwan 112, R.O.C.
JOURNAL   Location/Qualifiers
FEATURES   source
           1. .626
           /organism="SARS coronavirus TW-GD5"
           /mol_type="genomic RNA"
           /isolate="TW-GD5_SCI8"
           /db_xref="taxon:264382"
           <1. .>626
           /codon_start=1
           /product="replicase 1B"
           /protein_id="AA544825.1"
           /db_xref="GI:42741351"
           /translations="ISMATNYDLSVNNARLAKHYVYIGDPAQLPAPRTLLTKGTLEP
           EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCQKMFYK
           GVITHDVSSAINRPOIGVREFLTRNPARKAVFISPYNSONAVASKILGLPTQTVD
           SQSEYDYVIFTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN
Query Match      100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db      427 CAGAACGCTGTAGCTTCAAAAATCT 451

RESULT 13
AY451940
LOCUS    SARS coronavirus TW-YM1 isolate TW-YM1_SCI8 replicase 1B gene,
partial cds.
ACCESSION AY451940.1 GI:42741352
VERSION   SARS coronavirus TW-YM1
KEYWORDS  Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
ORGANISM  1 (bases 1 to 626)
           Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Molecular Epidemiology of SARS in Taiwan - Development of a Simple
           Method for Tracing the Origin and Dissemination of SARS
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS   Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Direct Submission
TITLE     Submitted (28-OCT-2003) National Yang-Ming University, AIDS
           Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
           Taipei, Taiwan 112, R.O.C.
JOURNAL   Location/Qualifiers
FEATURES   source
           1. .626
           /organism="SARS coronavirus TW-GD4"
           /mol_type="genomic RNA"
           /isolate="TW-GD4_SCI8"
           /db_xref="taxon:264381"
           <1. .>626
           /codon_start=1
           /product="replicase 1B"
           /protein_id="AA544824.1"
           /db_xref="GI:42741349"
           /translations="ISMATNYDLSVNNARLAKHYVYIGDPAQLPAPRTLLTKGTLEP
           EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCQKMFYK
           GVITHDVSSAINRPOIGVREFLTRNPARKAVFISPYNSONAVASKILGLPTQTVD
           SQSEYDYVIFTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN
Query Match      100.0%; Score 25; DB 14; Length 626;

```

```

Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db      427 CAGAACGCTGTAGCTTCAAAAATCT 451

RESULT 12
AY451939
LOCUS    SARS coronavirus TW-GD5 isolate TW-GD5_SCI8 replicase 1B gene,
partial cds.
ACCESSION AY451939.1 GI:42741350
VERSION   SARS coronavirus TW-GD5
KEYWORDS  Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
ORGANISM  1 (bases 1 to 626)
           Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Molecular Epidemiology of SARS in Taiwan - Development of a Simple
           Method for Tracing the Origin and Dissemination of SARS
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS   Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Direct Submission
TITLE     Submitted (28-OCT-2003) National Yang-Ming University, AIDS
           Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
           Taipei, Taiwan 112, R.O.C.
JOURNAL   Location/Qualifiers
FEATURES   source
           1. .626
           /organism="SARS coronavirus TW-GD5"
           /mol_type="genomic RNA"
           /isolate="TW-GD5_SCI8"
           /db_xref="taxon:264382"
           <1. .>626
           /codon_start=1
           /product="replicase 1B"
           /protein_id="AA544825.1"
           /db_xref="GI:42741351"
           /translations="ISMATNYDLSVNNARLAKHYVYIGDPAQLPAPRTLLTKGTLEP
           EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCQKMFYK
           GVITHDVSSAINRPOIGVREFLTRNPARKAVFISPYNSONAVASKILGLPTQTVD
           SQSEYDYVIFTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN
Query Match      100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db      427 CAGAACGCTGTAGCTTCAAAAATCT 451

RESULT 13
AY451940
LOCUS    SARS coronavirus TW-YM1 isolate TW-YM1_SCI8 replicase 1B gene,
partial cds.
ACCESSION AY451940.1 GI:42741352
VERSION   SARS coronavirus TW-YM1
KEYWORDS  Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
ORGANISM  1 (bases 1 to 626)
           Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Molecular Epidemiology of SARS in Taiwan - Development of a Simple
           Method for Tracing the Origin and Dissemination of SARS
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS   Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
           Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
           Direct Submission
TITLE     Submitted (28-OCT-2003) National Yang-Ming University, AIDS
           Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
           Taipei, Taiwan 112, R.O.C.
JOURNAL   Location/Qualifiers
FEATURES   source
           1. .626
           /organism="SARS coronavirus TW-GD4"
           /mol_type="genomic RNA"
           /isolate="TW-GD4_SCI8"
           /db_xref="taxon:264381"
           <1. .>626
           /codon_start=1
           /product="replicase 1B"
           /protein_id="AA544824.1"
           /db_xref="GI:42741349"
           /translations="ISMATNYDLSVNNARLAKHYVYIGDPAQLPAPRTLLTKGTLEP
           EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCQKMFYK
           GVITHDVSSAINRPOIGVREFLTRNPARKAVFISPYNSONAVASKILGLPTQTVD
           SQSEYDYVIFTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN
Query Match      100.0%; Score 25; DB 14; Length 626;

```

TITLE Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Molecular Epidemiology of SARS in Taiwan - Development of a Simple Method for Tracing the Origin and Dissemination of SARS

## JOURNAL

## REFERENCE

AUTHORS 2 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,  
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.

## TITLE

JOURNAL Direct Submission  
Submitted (28-OCT-2003) National Yang-Ming University, AIDS  
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,  
Taipei, Taiwan 112, R.O.C.

## FEATURES

source Location/Qualifiers

1. 626  
/organism="SARS coronavirus TW-YM1"  
/mol\_type="genomic RNA"  
/isolate="TW-YM1\_SC18"  
/db\_xref="taxon:264383"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AAS44826.1"  
/db\_xref="GI:42741353"

## CDS

/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLTLTKGTLEP  
EYFNSVCLMKITIGDPMFLGTCRCPCPAEIVDTVSALVDNKLKAHKDKSAOCFKMFYK  
GVITHDVSSAINRPOIGVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVD  
SOGSEYDYVIFTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLDY"

## ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 0.53;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25  
|||||

Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451  
|||||

## RESULT 14

## AY451941

## LOCUS

DEFINITION SARS coronavirus TW-YM2 isolate TW-YM2\_SC18 replicase 1B gene,  
partial cds.

## ACCESSION

## AY451941

## VERSION

## AY451941.1

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLTLTKGTLEP  
EYFNSVCLMKITIGDPMFLGTCRCPCPAEIVDTVSALVDNKLKAHKDKSAOCFKMFYK  
GVITHDVSSAINRPOIGVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVD  
SOGSEYDYVIFTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLDY"

## ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 0.53;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25  
|||||

Db 427 CAGAACGCTGTAGCTTCAAAAATCT 451  
|||||

## RESULT 15

## AY451942

## LOCUS

DEFINITION SARS coronavirus TW-YM3 isolate TW-YM3\_SC18 replicase 1B gene,  
partial cds.

## ACCESSION

## AY451942

## VERSION

## AY451942.1

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

Search completed: May 16, 2005, 02:00:23

Job time : 598.656 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 17:17:00 ; Search time 320.703 Seconds  
(without alignments)  
461.466 Million cell updates/sec

Title: US-10-808-187A-2471

Perfect score: 25  
Sequence: 1 cagaacgctgtagcttcaaaaatct 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseqn16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	29751	12	ADJ39000 SARS coro
2	18.6	74.4	493	4	ABa58348 Human foe
3	18.6	74.4	493	4	Aai37976 Probe #66
4	18.6	74.4	493	4	Aak32125 Human bon
5	18.6	74.4	493	4	Aak06447 Human bra
6	18.6	74.4	493	4	Abs31821 Human liv
7	18.6	74.4	493	6	ABs06892 Human gen
8	18.2	72.8	1056	13	ADR26577 Breast ca
9	17.8	71.2	1524	8	ACA20744 Prokaryot
10	17.8	71.2	1569	9	ADA32694 DNA encod
11	17.8	71.2	83836	11	ACN45080 Mouse gen
12	17.6	70.4	60	6	ABT12027 E coli ex
13	17.6	70.4	487	9	ACH26262 Human adu
14	17.6	70.4	487	9	ACH27372 Human adu
15	17.6	70.4	562	13	ACN50445 Cotton ma
16	17.6	70.4	600	13	ADT05557 Haemophil
17	17.6	70.4	609	6	ABN64132 Human can
18	17.6	70.4	712	12	AD42255 Plant tra
19	17.6	70.4	712	12	ADO02732 Rice orth
20	17.6	70.4	712	12	ADO62314 Transcrip

C 21	17.6	70.4	972	4	AAS53300	Aas53300 Haemophil
C 22	17.6	70.4	972	8	ACA34137	Aca34137 Prokaryot
C 23	17.6	70.4	972	13	ADT05583	Adt05583 Haemophil
C 24	17.6	70.4	1040	6	ABT11987	Abt11987 E coli ex
C 25	17.6	70.4	2183	4	AAH77024	Aah77024 Human hel
C 26	17.6	70.4	2183	6	ABA95848	Abag5848 Human hel
C 27	17.6	70.4	2687	12	ADJ39475	Adj39475 Plant CDN
C 28	17.6	70.4	2735	4	AAF32744	Aaf32744 Human sec
C 29	17.6	70.4	2878	4	AAS02416	Aas02416 Human sec
C 30	17.6	70.4	2878	6	ABL90683	Ab190683 Human pol
C 31	17.6	70.4	2878	8	ABZ73672	Abz73672 Secreted
C 32	17.6	70.4	2878	10	ABZ67269	Abz67269 Human sec
C 33	17.6	70.4	2937	6	ABV78121	Abv78121 Human eph
C 34	17.6	70.4	2997	6	ABZ35697	Abz35697 Human eph
C 35	17.6	70.4	2997	6	ABX09940	Abx09940 Human eph
C 36	17.6	70.4	2997	6	ABL91662	Ab191662 Human pol
C 37	17.6	70.4	3832	13	ADT05409	Adt05409 Haemophil
C 38	17.6	70.4	4283	4	AAS25994	Aas25994 Human CDN
C 39	17.6	70.4	4283	8	ABX73335	Abx73335 Human nov
C 40	17.6	70.4	4523	8	ABZ34853	Abz34853 Coding se
C 41	17.6	70.4	4529	2	AAT02949	Aat02949 EPH-like
C 42	17.6	70.4	4569	4	AAI60359	Aai60359 Human pol
C 43	17.6	70.4	4569	4	AAI60358	Aai60358 Human pol
C 44	17.6	70.4	4569	4	AAI60360	Aai60360 Human pol
C 45	17.6	70.4	4774	12	ADM86923	Adm86923 Human pro

ALIGNMENTS

RESULT 1

ADJ39000

ID ADJ39000 standard; DNA; 29751 BP.

XX

AC ADJ39000;

XX

DT 06-MAY-2004 (first entry)

XX

DE SARS coronavirus nucleotide sequence.

XX

KW small interfering RNA; siRNA; modified ribonucleotide;

KW viral replication inhibition; hepatitis C virus; HCV; hepatitis C;

KW antiinflammatory; hepatotropic; virucide; hepatitis A virus;

KW hepatitis D virus; hepatitis E virus; Ebola virus; influenza virus;

KW rotavirus; reovirus; retrovirus; poliovirus; human papilloma virus;

KW metapneumoniavirus; coronavirus; viral infection; gene; ds.

XX

OS SARS coronavirus.

XX

PN WO2004011647-A1.

XX

PD 05-FEB-2004.

XX

PF 25-JUL-2003; 2003WO-US023104.

XX

PR 26-JUL-2002; 2002US-0398605P.

 XX || PA | (CHIR ) CHIRON CORP. |
XX	
PI	Han J, Seo MY, Houghton M;
XX	
DR	WPI; 2004-143862/14.
XX	
PT	New RNase resistant small interfering RNA, useful for treating viral
XX	infections, e.g., hepatitis C, influenza virus or coronavirus infection.
XX	Example 10; Fig 3; 74pp; English.
CC	The present invention describes a small interfering RNA (siRNA) which
XX	comprises a modified ribonucleotide, where the siRNA is resistant to
CC	RNase and retains the ability to inhibit viral replication. Also
XX	described: (1) inactivating a virus in a patient; (2) making a modified
CC	siRNA that targets a nucleic acid sequence in a virus; (3) a double-

CC stranded RNA molecule of 10-30 nucleotides that inhibits replication of  
 CC hepatitis C virus (HCV); (4) inducing targeted RNA interference toward  
 CC HCV in hepatic cells; (5) inhibiting replication of HCV; (6) a vector  
 CC comprising a DNA segment encoding the RNA molecule; (7) a host cell  
 CC comprising the vector of (6); (8) inhibiting replication of HCV in cells  
 CC carrying HCV; (9) treating hepatitis C in a subject; (10) a modified  
 CC siRNA molecule comprising a double-stranded RNA molecule of 10-30  
 CC nucleotides in length, which mediates RNA interference toward a target  
 CC agent or virus and is linked to at least one receptor-binding ligand; and  
 CC (11) inducing targeted RNA interference in a patient. The modified siRNA  
 CC molecules have antiinflammatory, hepatotropic and virucide activities.  
 CC The modified RNA molecules are useful for inactivating virus in mammalian  
 CC cells. The siRNAs are useful for treating hepatitis C virus, hepatitis A  
 CC virus, hepatitis D virus, hepatitis E virus, Ebola virus, influenza  
 CC virus, rotavirus, reovirus, coronavirus, poliovirus, human papilloma  
 CC virus, metapneumovirus or coronavirus infections. The methods of the  
 CC invention can be used to correct or compensate for cellular physiological  
 CC abnormalities involved in conferring susceptibility to viral infections  
 CC in patients and/or alleviate symptoms of a viral infection in patients.  
 CC The present sequence represents the SARS coronavirus nucleotide sequence,  
 CC which is used in an example from the present invention.

XX Sequence 29751 BP; 8478 A; 5941 C; 6188 G; 9144 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 12; Length 29751;  
 Best Local Similarity 100.0%; Pred. No. 0.19;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25  
 |||||  
 Db 17718 CAGAACGCTGTAGCTTCAAAAATCT 17742

## RESULT 2

ABA58348  
 ID ABA58348 standard; DNA; 493 BP.

XX ABA58348;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #6653.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human foetal liver.

XX Claim 1; SEQ ID NO 6653; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring

CC human gene expression in a sample derived from human foetal liver. The  
 CC single exon nucleic acid probes may be used for predicting, measuring and  
 CC displaying gene expression in samples derived from human foetal liver. The  
 CC present sequence is a single exon nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 493 BP; 182 A; 71 C; 70 G; 170 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 4; Length 493;  
 Best Local Similarity 84.0%; Pred. No. 1.1e+02;  
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25  
 |||||  
 Db 150 CAGAACACAGTATCTTCCAAAATCT 174

## RESULT 3

AAI37976  
 ID AAI37976 standard; DNA; 493 BP.

XX AAI37976;

XX 17-OCT-2001 (first entry)

XX Probe #6662 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder; ss.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.

XX Claim 25; SEQ ID NO 6662; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP).

XX The present sequence is one such probe. The probes are useful for

XX producing a microarray for predicting, measuring and displaying gene

XX expression in samples derived from human placenta. The probes are useful

XX for antenatal diagnosis of human genetic disorders

XX Sequence 493 BP; 182 A; 71 C; 70 G; 170 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 4; Length 493;  
 Best Local Similarity 84.0%; Pred. No. 1.1e+02;  
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25  
 |||||

KW	ss.
XX	
XX	Homo sapiens.
OS	
XX	
PN	W0200157275-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30--JAN-2001; 2001W0-US0000667.
XX	
PR	04-FEB-2000; 2000US-0180312P.
PR	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.
PR	03-AUG-2000; 2000US-00632366.
PR	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.
PR	04-OCT-2000; 2000GB-00024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI; 2001-483446/52.
XX	
PT	Single exon nucleic acid probes for analyzing gene expression in human brains.
PT	
XX	
PS	Example 4; SEQ ID NO 6438; 650pp + Sequence Listing; English.
XX	
CC	The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention
XX	
SQ	Sequence 493 BP; 182 A; 71 C; 70 G; 170 T; 0 U; 0 Other;
	Query Match 74.4%; Score 18.6; DB 4; Length 493;
	Best Local Similarity 84.0%; Pred.No.1.1e-02;
	Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps
Qy	1 CAGAACGCTGTAGCTTCAAAAATCT 25
	-
Db	150 CAGAACACAGTATCTTCCAAAATCT 174
RESULT 6	
ABS31821	
ID	ABS31821 standard; DNA; 493 BP.
XX	
AC	ABS31821;
XX	
DT	25-FEB-2003 (first entry)
XX	
DE	Human liver single exon probe, SEQ ID No 6811.
XX	
KW	Human; single exon nucleic acid probe; liver; cirrhosis;
KW	hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW	coronary heart disease; ss.
XX	
OS	Homo sapiens.
XX	
FN	W0200157273-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001W0-US0000664.
XX	
PR	04-FEB-2000; 2000US-0180312P.
PR	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.

```

PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488998/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 1; SEQ ID NO 6811; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABS25011-ABS51005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 493 BP; 182 A; 71 C; 70 G; 170 T; 0 U; 0 Other;
Query Match 74.4%; Score 18.6; DB 4; Length 493;
Best Local Similarity 84.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db 150 CAGAACACAGTATCTTCCAAAATCT 174
RESULT 7
ID ABS06892
AC ABS06892 standard; DNA; 493 BP.
AC
XX ABS06892;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon probe from lung SEQ ID No 6883.
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US0000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.

```

```

PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
XX Claim 1; SEQ ID NO 6883; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX; the novel set of probes which hybridise at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of the
XX array; identifying exons in a eukaryotic genome, comprising (a)
XX algorithmically predicting at least one exon from genomic sequences of
XX the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene expression
XX analysis, and for identifying exons in a gene, particularly using human
XX lung derived mRNA and for the study of lung diseases such as asthma, lung
XX cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
XX disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
XX tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
XX Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
XX histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
XX Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
XX dyskinesia, pulmonary hypertension and hyaline membrane disease. The
XX present sequence is a single exon probe of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 493 BP; 182 A; 71 C; 70 G; 170 T; 0 U; 0 Other;
XX
XX Query Match 74.4%; Score 18.6; DB 6; Length 493;
XX Best Local Similarity 84.0%; Pred. No. 1.1e+02;
XX Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
Db 150 CAGAACACAGTATCTTCCAAAATCT 174
RESULT 8
ID ADR26577
AC ADR26577 standard; DNA; 1056 BP.
XX
XX ADR26577;
XX
XX 21-OCT-2004 (first entry)

```

```
XX DE Breast cancer prognosis marker #2438.
XX KW ds; breast cancer; prognosis; gene expression; diagnosis.
XX OS Homo sapiens.
XX PN WO2004065545-A2.
XX PD 05-AUG-2004.
XX PF 15-JAN-2004; 2004WO-US001100.
XX PP 15-JAN-2003; 2003US-00342887.
XX PR (ROSE-) ROSETTA INPHARMATICS LLC.
XX PA (NECA-) NETHERLANDS CANCER INST.
XX PI Van't Veer LJ, He Y;
XX PF WIPI; 2004-593473/57.
XX PT Classifying a breast cancer patient according to prognosis comprises
XX PT determining the similarity between the level of expression of each of
XX PT five genes in a cell sample taken from patient, to control levels.
XX PS Disclosure; SEQ ID NO 2438; 226pp; English.
XX CC The invention relates to a method of classifying a breast cancer patient
XX CC according to prognosis by determining the similarity between the level of
XX CC expression of each of five genes for which markers are listed in the
XX CC specification, in a cell sample taken from the breast cancer patient, to
XX CC control levels of expression for each respective five genes to obtain a
XX CC patient similarity value. The methods are useful for classifying a breast
XX CC cancer patient according to prognosis. Kits and computer program products
XX CC are useful for data analysis using the diagnostic, prognostic and
XX CC statistical methods of the invention. This sequence corresponds to a
XX CC marker used in the method of the invention.
XX SQ Sequence 1056 BP; 389 A; 168 C; 167 G; 332 T; 0 U; 0 Other;

Query Match 72.8%; Score 18.2; DB 13; Length 1056;
Best Local Similarity 87.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 GAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 945 GAACGCTGTACCTTAAAAATTT 967
|||||

RESULT 9
ACA20744
ID ACA20744 standard; DNA; 1524 BP.
XX AC ACA20744;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #2401.
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX OS Acinetobacter baumannii.
XX PN WO20027183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US0009107.
XX PP 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
```

```
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PF Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PF Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR P-PSDB; ABUI6874.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 8614; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX CC prokaryotic essential genes. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1524 BP; 452 A; 320 C; 341 G; 411 T; 0 U; 0 Other;

Query Match 71.2%; Score 17.8; DB 8; Length 1524;
Best Local Similarity 90.5%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 ACGCTGTAGCTTCAAAAATCT 25
|||||
Db 1457 ACGCTGTAGCTTCAAAAATCT 1477
|||||

RESULT 10
ADA32694
ID ADA32694 standard; DNA; 1569 BP.
XX AC ADA32694;
XX DT 20-NOV-2003 (first entry)
XX DE DNA encoding Acinetobacter baumannii protein #3981.
XX KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
XX KW vaccine; plant biocontrol agent.
```

```

XX OS Acinetobacter baumannii.
XX PN US6562958-B1.
XX PD 13-MAY-2003.
XX PF 04-JUN-1999; 99US-00328352.
XX PR 09-JUN-1998; 98US-0088701P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton G, Bush D;
XX WPI; 2003-576092/54.
XX DR P-PSDB; ADA36820.
XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX PT for diagnosing a bacterial disease, as components of antibacterial
XX PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX PT plants.
XX PS Example; SEQ ID NO 3981; 328pp; English.
XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX CC The A. baumannii nucleic acids and polypeptides are useful as reagents
XX CC for diagnosing a bacterial disease, as components of antibacterial
XX CC vaccines, as targets for antibacterial drugs, to detect the presence of
XX CC A. baumannii and other Acinetobacter species in a sample, in screening
XX CC compounds for the ability to interfere with the A. baumannii life cycle
XX CC or to inhibit A. baumannii infection, and as biocontrol agents for
XX CC plants. The present sequence represents DNA encoding an A. baumannii
XX CC protein.
XX SQ Sequence 1569 BP; 461 A; 331 C; 352 G; 425 T; 0 U; 0 Other;

Query Match 71.2%; Score 17.8; DB 9; Length 1569;
Best Local Similarity 90.5%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 ACGCTGTAGCTTCAAAAATCT 25
DB 1499 ACGCAGTAGCTCCAAAATCT 1519

RESULT 11
ACN45080/c
ID ACN45080 standard; DNA; 83836 BP.
XX AC ACN45080;
XX DT 18-NOV-2004 (first entry)
XX DE Mouse genomic sequence MCG7826.
XX KW Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX OS Mus musculus.
XX PN WO2003073826-A2.
XX PD 12-SEP-2003.
XX PF 28-FEB-2003; 2003WO-US006235.
XX PR 01-MAR-2002; 2002US-00087192.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX WPI; 2003-328604/31.

```

```

XX PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX PT comprises a nucleotide sequence.
XX PS Claim 1; SEQ ID NO 1849; Opp; English.
XX CC The present invention relates to novel DNA and protein sequences which
XX CC are associated with carcinomas. The sequences are useful for: (i) for
XX CC screening drug candidates; (ii) for screening of bioactive agent capable
XX CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX CC a bioactive agent capable of modulating the activity of CAP; (iv) for
XX CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX CC determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX CC carcinoma including lymphoma. The present sequence is one such CA coding
XX CC sequence. Note: This patent is an equivalent to basic patent
XX CC US2002182586A1, for which no sequence data was published
XX SQ Sequence 83836 BP; 23536 A; 15199 C; 15996 G; 26941 T; 0 U; 2164 Other;

Query Match 71.2%; Score 17.8; DB 11; Length 83836;
Best Local Similarity 90.5%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAAGCGTGTAGCTTCAAAAA 22
DB 16018 AGAACCCCTGTAGCTTAAAAA 15998

RESULT 12
AET12027/c
ID AET12027 standard; DNA; 60 BP.
XX AC AET12027;
XX DT 19-DEC-2002 (first entry)
XX DE E coli expressible CpG deprived gene related oligo SEQ ID No 48.
XX KW CpG; Escherichia coli; expression; DNA vaccine; gene therapy; ds.
XX OS Unidentified.
XX PN FR2821855-A1.
XX PD 13-SEP-2002.
XX PF 09-MAR-2001; 2001FR-00003274.
XX PR 09-MAR-2001; 2001FR-00003274.
XX PA (CAYL-) CAYLA SARL.
XX DR WPI; 2002-715203/78.
XX PT Preparing gene with no cytosine methylation, useful e.g. for expressing
XX PT proteins in Escherichia coli, by selecting codons to avoid presence of
XX PT CpG dinucleotides.
XX PS Disclosure; Fig 4; 152pp; French.
XX CC The invention relates to a method for preparing a gene that lacks CpG but
XX CC is still expressible in Escherichia coli by synthesis of a polynucleotide
XX CC sequence, encoding a protein expressible in E. coli but lacking the
XX CC codons NCG, CGN or NTA (N = any of A, C, G or T) or any codons that end
XX CC in C when the following codon starts with G. The CpG lacking genes are
XX CC especially useful as selection (antibiotic resistance) or marker (lacZ)
XX CC genes, but they (also CpG-free promoters, transcription units, origins of
XX CC replication, plasmids and cells for replication of the plasmids) can be
XX CC used for transformation, for biotechnological or medical purposes, e.g.

```



CC they are particularly well suited for preparation of DNA vaccines for use  
CC in humans or animals. The polynucleotides of the invention can be used  
CC for treating disorders by gene therapy. This polynucleotide sequence  
CC represents an oligonucleotide relating to the Escherichia coli  
CC expressible CpG gene of the invention  
XX  
SQ Sequence 60 BP; 19 A; 13 C; 14 G; 14 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 6; Length 60;  
Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATC 24  
||||| ||||| ||||| ||||| |||||  
Db 47 CAGAACATTGTGTTCCAAAATC 24

## RESULT 13

ACH26262  
ID ACH26262 standard; cDNA; 487 BP.

XX AC ACH26262;

XX AC ACH26262;

DT 13-OCT-2003 (first entry)

DE Human adult ovary cDNA #4642.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

KW genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX Claim 1; SEQ ID NO 13474; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC obtained in electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
SQ Sequence 487 BP; 123 A; 107 C; 139 G; 114 T; 0 U; 4 Other;  
Query Match 70.4%; Score 17.6; DB 9; Length 487;  
Best Local Similarity 83.3%; Pred. No. 3.3e+02;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATC 24  
||||| ||||| ||||| ||||| |||||  
Db 281 CAGCAGGCTGTAGCTCCAAAATC 304

## RESULT 14

ACH27372  
ID ACH27372 standard; cDNA; 487 BP.

XX AC ACH27372;

XX AC ACH27372;

DT 13-OCT-2003 (first entry)

DE Human adult ovary cDNA #5752.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

KW genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX Claim 1; SEQ ID NO 14584; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC obtained in electronic format directly from USPTO at

XX seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
SQ Sequence 487 BP; 122 A; 105 C; 139 G; 117 T; 0 U; 4 Other;



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:11:05 ; Search time 93.5547 Seconds  
(without alignments)  
437.251 Million cell updates/sec

Title: US-10-808-187A-2471  
Perfect score: 25  
Sequence: 1 cagaacgctgtagcttcaaaatct 25

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.2	72.8	84296	4	US-09-949-016-17375
C 2	17.8	71.2	1569	4	US-09-328-352-3981
C 3	17.6	70.4	4522	4	US-09-949-016-4030
C 4	17.6	70.4	4523	4	US-09-949-016-630
C 5	17.6	70.4	4529	2	US-08-449-645A-16
C 6	17.6	70.4	4529	2	US-08-702-367A-16
C 7	17.6	70.4	4529	5	PCT-US95-04681-16
C 8	17.6	70.4	6223	4	US-09-620-312D-459
C 9	17.6	70.4	6304	4	US-09-620-312D-461
C 10	17.6	70.4	6382	4	US-09-620-312D-460
C 11	17.6	70.4	13206	3	US-08-961-527-33
C 12	17.6	70.4	1830121	4	US-09-557-884-1
C 13	17.6	70.4	1830121	4	US-09-643-990A-1
C 14	17.2	68.8	43690	4	US-09-949-016-13904
C 15	17.2	68.8	1230025	4	US-09-198-452A-1
C 16	17.2	68.8	1230230	4	US-09-438-185A-1
C 17	17	68.0	493	4	US-09-270-767-9570
C 18	17	68.0	493	4	US-09-270-767-24852
C 19	17	68.0	15027	4	US-09-949-016-12660
C 20	17	68.0	15036	4	US-09-949-016-13351
C 21	17	68.0	101300	4	US-09-949-016-16108
C 22	17	68.0	580073	4	US-08-545-528D-1
C 23	16.6	66.4	601	4	US-09-949-016-17626
C 24	16.6	66.4	601	4	US-09-949-016-22400
C 25	16.6	66.4	601	4	US-09-949-016-22401
C 26	16.6	66.4	601	4	US-09-949-016-26775
C 27	16.6	66.4	601	4	US-09-949-016-26776

28	16.6	66.4	601	4	US-09-949-016-26777	Sequence 26777, A
29	16.6	66.4	601	4	US-09-949-016-26778	Sequence 26778, A
C 30	16.6	66.4	601	4	US-09-949-016-31117	Sequence 31117, A
31	16.6	66.4	601	4	US-09-949-016-31155	Sequence 31155, A
32	16.6	66.4	601	4	US-09-949-016-36130	Sequence 36130, A
33	16.6	66.4	601	4	US-09-949-016-54598	Sequence 54598, A
34	16.6	66.4	601	4	US-09-949-016-54630	Sequence 54630, A
35	16.6	66.4	601	4	US-09-949-016-54662	Sequence 54662, A
36	16.6	66.4	601	4	US-09-949-016-61461	Sequence 61461, A
37	16.6	66.4	601	4	US-09-949-016-119027	Sequence 119027, A
38	16.6	66.4	601	4	US-09-949-016-119063	Sequence 119063, A
39	16.6	66.4	601	4	US-09-949-016-119099	Sequence 119099, A
40	16.6	66.4	601	4	US-09-949-016-119135	Sequence 119135, A
41	16.6	66.4	601	4	US-09-949-016-119171	Sequence 119171, A
42	16.6	66.4	601	4	US-09-949-016-119207	Sequence 119207, A
C 43	16.6	66.4	601	4	US-09-949-016-120285	Sequence 120285, A
C 44	16.6	66.4	601	4	US-09-949-016-142039	Sequence 142039, A
45	16.6	66.4	601	4	US-09-949-016-154205	Sequence 154205, A

ALIGNMENTS

RESULT 1  
US-09-949-016-17375/c  
; Sequence 17375, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17375  
; LENGTH: 84296  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(84296)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17375

Query Match 72.8%; Score 18.2; DB 4; Length 84296;  
Best Local Similarity 87.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GAACGCTGTAGCTTCAAAATCT 25  
|||||  
Db 82938 GAACGCTGTAGCTTCAAAATCT 82916

RESULT 2

US-09-328-352-3981  
; Sequence 3981, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252

```
; SEQ ID NO 3981
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3981

Query Match          71.2%; Score 17.8; DB 4; Length 1569;
Best Local Similarity 90.5%; Pred. No. 56;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 ACGCTGTAGCTTCAAAAATCT 25
    ||| ||||| ||||| ||||| |||||
Db 1499 AGCAGTAGCTTCAAAAATCT 1519

RESULT 3
US-09-949-016-4030/c
; Sequence 4030, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4030
; LENGTH: 4522
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4030

Query Match          70.4%; Score 17.6; DB 4; Length 4522;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGAACGCTGTAGCTTCAAAAATCT 25
    ||| ||||| ||||| ||||| |||||
Db 1830 AGACAGCTGTAGCTTCAAAACATT 1807

RESULT 4
US-09-949-016-630/c
; Sequence 630, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 630
; LENGTH: 4523
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-630
```

```
Query Match          70.4%; Score 17.6; DB 4; Length 4523;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGAACGCTGTAGCTTCAAAAATCT 25
    ||| ||||| ||||| ||||| |||||
Db 1830 AGACAGCTGTAGCTTCAAAACATT 1807

RESULT 5
US-08-449-645A-16/c
; Sequence 16, Application US/08449645A
; Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,645A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4529 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 186..3182
US-08-449-645A-16

Query Match          70.4%; Score 17.6; DB 2; Length 4529;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGAACGCTGTAGCTTCAAAAATCT 25
    ||| ||||| ||||| ||||| |||||
Db 1830 AGACAGCTGTAGCTTCAAAACATT 1807

RESULT 6
US-08-702-367A-16/c
; Sequence 16, Application US/08702367A
; Patent No. 5981246
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
```

;; COUNTRY: USA  
;; ZIP: 91320  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/702,367A  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Winter, Robert B.  
;; REFERENCE/DOCKET NUMBER: A-287  
;; INFORMATION FOR SEQ ID NO: 16:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4529 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 186..3182  
;; US-08-702-367A-16

Query Match 70.4%; Score 17.6; DB 2; Length 4529;  
Best Local Similarity 83.3%; Pred. No. 88;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGACGCTGTAGCTTCAAAATCT 25  
DB 1830 AGACGCTGTAGCTTCAAAATTT 1807

RESULT 7  
PCT-US95-04681-16/c  
;; GENERAL INFORMATION:  
;; APPLICANT: Fox, Gary M.  
;; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine  
;; TITLE OF INVENTION: Kinases  
;; NUMBER OF SEQUENCES: 28  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Amgen Patent Operations/RBW  
;; STREET: 1840 Dehavilland Drive  
;; CITY: Thousand Oaks  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 91320  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/04681  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Winter, Robert B.  
;; REFERENCE/DOCKET NUMBER: A-287  
;; INFORMATION FOR SEQ ID NO: 16:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4529 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 186..3182  
;; PCT-US95-04681-16

Query Match 70.4%; Score 17.6; DB 5; Length 4529;  
Best Local Similarity 83.3%; Pred. No. 88;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 2 AGACGCTGTAGCTTCAAAATCT 25  
DB 1830 AGACGCTGTAGCTTCAAAATTT 1807

RESULT 8  
US-09-620-312D-459  
;; Sequence 459, Application US/09620312D  
;; Patent No. 659662  
;; GENERAL INFORMATION:  
;; APPLICANT: Tang, Y. Tom  
;; APPLICANT: Liu, Chenghua  
;; APPLICANT: Asundi, Vinod  
;; APPLICANT: Zhang, Jie  
;; APPLICANT: Ren, Feiyan  
;; APPLICANT: Chen, Rui-hong  
;; APPLICANT: Zhao, Qing A.  
;; APPLICANT: Wehrman, Tom  
;; APPLICANT: Xue, Aidong J.  
;; APPLICANT: Yang, Yonghong  
;; APPLICANT: Wang, Jian-Rui  
;; APPLICANT: Zhou, Ping  
;; APPLICANT: Ma, Yunging  
;; APPLICANT: Wang, Dunrui  
;; APPLICANT: Wang, Zhiwei  
;; APPLICANT: John Tillinghast  
;; APPLICANT: Drmanac, Radoje T.  
;; TITLE OF INVENTION: No. 6589662el Nucleic Acids and  
;; TITLE OF INVENTION: Polypeptides  
;; FILE REFERENCE: 784CIP28  
;; CURRENT APPLICATION NUMBER: US/09/620,312D  
;; CURRENT FILING DATE: 2000-07-19  
;; PRIOR APPLICATION NUMBER: 09/552,317  
;; PRIOR FILING DATE: 2000-04-25  
;; PRIOR APPLICATION NUMBER: 09/488,725  
;; PRIOR FILING DATE: 2000-01-21  
;; NUMBER OF SEQ ID NOS: 1105  
;; SOFTWARE: pt\_FL\_genes Version 1.0  
;; SEQ ID NO 459  
;; LENGTH: 6223  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (59)..(4132)  
;; US-09-620-312D-459

Query Match 70.4%; Score 17.6; DB 4; Length 6223;  
Best Local Similarity 83.3%; Pred. No. 94;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGACGCTGTAGCTTCAAAATC 24  
DB 3660 CAGACGCTGTAGCTTCAAAATTC 3683

RESULT 9  
US-09-620-312D-461  
;; Sequence 461, Application US/09620312D  
;; Patent No. 659662  
;; GENERAL INFORMATION:  
;; APPLICANT: Tang, Y. Tom  
;; APPLICANT: Liu, Chenghua  
;; APPLICANT: Asundi, Vinod  
;; APPLICANT: Zhang, Jie  
;; APPLICANT: Ren, Feiyan  
;; APPLICANT: Chen, Rui-hong  
;; APPLICANT: Zhao, Qing A.



```
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 70.4%; Score 17.6; DB 4; Length 1830121;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 519422 AGAACGCTGTTCATCAAAAATCT 519399

RESULT 13
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
;
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
```

```
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 70.4%; Score 17.6; DB 4; Length 1830121;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 519422 AGAACGCTGTTCATCAAAAATCT 519399

RESULT 14
US-09-949-016-13904
; Sequence 13904, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13904
; LENGTH: 43690
; TYPE: DNA
; ORGANISM: Human
;
; US-09-949-016-13904

Query Match 68.8%; Score 17.2; DB 4; Length 43690;
Best Local Similarity 86.4%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGAACGCTGTAGCTTCAAAAAT 23
|||||
Db 30403 AGAACGCTGTAGCTGAAAAAT 30424

RESULT 15
US-09-198-452A-1
; Sequence 1, Application US/09198452A
```





OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (660001)..(675000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (675001)..(690000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (690001)..(705000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (705001)..(720000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (720001)..(735000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (735001)..(750000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (750001)..(765000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (765001)..(780000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (780001)..(795000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (795001)..(810000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (810001)..(825000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (825001)..(840000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (840001)..(855000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (855001)..(870000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (870001)..(885000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (885001)..(900000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature  
LOCATION: (900001)..(915000)  
OTHER INFORMATION: n=a or c or g or t  
NAME/KEY: misc\_feature

Query Match 68.8%; Score 17.2; DB 4; Length 1230025;  
Best Local Similarity 86.4%; Pred. No. 3.8e+02;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AGAACGCTGTAGCTTCAAAAAT 23  
Db 228583 AGAACCTATAGATTCAAAAAT 228604

Search completed: May 16, 2005, 06:11:23  
Job time : 101.555 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 02:00:36 ; Search time 1185.16 Seconds

(without alignments)  
129.102 Million cell updates/sec

Title: US-10-808-187A-2471

Perfect score: 25

Sequence: 1 cagaacgctagcttcaaaaatct 25

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_5/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	25	100.0	25	19	US-10-808-187-2471
2	25	100.0	1213	19	US-10-808-187-11
3	25	100.0	5262	19	US-10-699-936-9
4	25	100.0	28920	19	US-10-889-447-5
5	25	100.0	28920	19	US-10-889-447-6
6	25	100.0	29291	19	US-10-889-447-4
7	25	100.0	29430	19	US-10-889-447-7
8	25	100.0	29727	18	US-10-839-729-15
9	25	100.0	29727	18	US-10-827-757-1
10	25	100.0	29727	19	US-10-889-447-8
11	25	100.0	29727	19	US-10-699-936-1
					Sequence 2471, Ap
					Sequence 11, Appl
					Sequence 9, Appl
					Sequence 5, Appl
					Sequence 6, Appl
					Sequence 4, Appl
					Sequence 7, Appl
					Sequence 15, Appl
					Sequence 1, Appl
					Sequence 8, Appl
					Sequence 1, Appl

12	25	100.0	29736	18	US-10-839-729-17	Sequence 17, Appl
13	25	100.0	29736	19	US-10-889-447-9	Sequence 9, Appl
14	25	100.0	29736	19	US-10-699-936-3	Sequence 3, Appl
15	25	100.0	29742	18	US-10-839-729-16	Sequence 16, Appl
16	25	100.0	29742	19	US-10-808-187-15	Sequence 15, Appl
17	25	100.0	29742	19	US-10-808-187-16	Sequence 16, Appl
18	25	100.0	29742	19	US-10-808-187-240	Sequence 240, Appl
19	25	100.0	29742	19	US-10-808-187-737	Sequence 737, Appl
20	25	100.0	29742	19	US-10-808-187-1108	Sequence 1108, Appl
21	25	100.0	29742	19	US-10-808-187-1590	Sequence 1590, Appl
22	25	100.0	29742	19	US-10-808-187-1965	Sequence 1965, Appl
23	25	100.0	29742	19	US-10-889-447-10	Sequence 10, Appl
24	25	100.0	29751	18	US-10-839-729-14	Sequence 14, Appl
25	25	100.0	29751	19	US-10-856-529-1	Sequence 1, Appl
26	25	100.0	29751	19	US-10-626-879-67	Sequence 67, Appl
27	25	100.0	29751	19	US-10-889-447-1	Sequence 1, Appl
28	25	100.0	29751	19	US-10-889-447-2	Sequence 2, Appl
29	25	100.0	29751	19	US-10-699-936-2	Sequence 2, Appl
30	23.4	93.6	646	19	US-10-808-187-1	Sequence 1, Appl
31	18.6	74.4	493	9	US-09-864-761-11173	Sequence 11173, A
32	18.2	72.8	1056	17	US-10-172-118-2438	Sequence 2438, Appl
33	18.2	72.8	1056	17	US-10-342-887-2438	Sequence 2438, Appl
34	17.8	71.2	1524	17	US-10-282-122A-8614	Sequence 8614, Appl
35	17.8	71.2	83836	13	US-10-087-192-1849	Sequence 1849, Appl
36	17.6	70.4	60	18	US-10-469-851-48	Sequence 48, Appl
37	17.6	70.4	161	9	US-09-783-590-8336	Sequence 8336, Appl
38	17.6	70.4	487	10	US-09-918-995-13474	Sequence 13474, A
39	17.6	70.4	487	10	US-09-918-995-14584	Sequence 14584, A
40	17.6	70.4	562	18	US-10-021-323-5226	Sequence 5226, Appl
41	17.6	70.4	663	13	US-10-027-632-245266	Sequence 245266, Appl
42	17.6	70.4	663	17	US-10-027-632-245266	Sequence 245266, Appl
43	17.6	70.4	712	17	US-10-374-780A-718	Sequence 718, Appl
44	17.6	70.4	712	17	US-10-412-699B-1145	Sequence 1145, Appl
45	17.6	70.4	934	17	US-10-424-599-83579	Sequence 83579, A

ALIGNMENTS

RESULT 1  
US-10-808-187-2471  
; Sequence 2471, Application US/10808187  
; Publication No. US2005009009A1  
; GENERAL INFORMATION:  
; APPLICANT: PEIRIS, JOSEPH S. M.  
; APPLICANT: YUEN, KWOK YUNG  
; APPLICANT: POON, LIT MAN  
; APPLICANT: GUAN, YI  
; APPLICANT: CHAN, KWOK HUNG  
; APPLICANT: NICHOLLS, JOHN  
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE  
; TITLE OF INVENTION: RESPIRATORY SYNDROME (SARS)  
; FILE REFERENCE: V9661.0078  
; CURRENT APPLICATION NUMBER: US/10/808,187  
; CURRENT FILING DATE: 2004-03-24  
; PRIOR APPLICATION NUMBER: 60/457,031  
; PRIOR FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: 60/457,730  
; PRIOR FILING DATE: 2003-03-26  
; PRIOR APPLICATION NUMBER: 60/459,931  
; PRIOR FILING DATE: 2003-04-02  
; PRIOR APPLICATION NUMBER: 60/460,357  
; PRIOR FILING DATE: 2003-04-03  
; PRIOR APPLICATION NUMBER: 60/461,265  
; PRIOR FILING DATE: 2003-04-08  
; PRIOR APPLICATION NUMBER: 60/462,805  
; PRIOR FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 60/468,139  
; PRIOR FILING DATE: 2003-05-05  
; PRIOR APPLICATION NUMBER: 60/464,886  
; PRIOR FILING DATE: 2003-04-23  
; PRIOR APPLICATION NUMBER: 60/471,200  
; PRIOR FILING DATE: 2003-05-16

```
; NUMBER OF SEQ ID NOS: 2476
; SOFTWARE: PatentIn ver. 3.2
; SEQ ID NO 2471
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-808-187-2471
```

```
Query Match      100.0%; Score 25; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.071; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;
```

```
QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 1 CAGAACGCTGTAGCTTCAAAAATCT 25
```

## RESULT 2

```
US-10-808-187-11
; Sequence 11, Application US/10808187
; Publication No. US2005000909A1
; GENERAL INFORMATION:
; APPLICANT: PEIRIS, JOSEPH S. M.
; APPLICANT: YUEN, KWOK YUNG
; APPLICANT: POON, LIT MAN
; APPLICANT: GUAN, YI
; APPLICANT: CHAN, KWOK HUNG
; APPLICANT: NICHOLLS, JOHN
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE
; FILE REFERENCE: V9661.0078
; CURRENT APPLICATION NUMBER: US/10/808,187
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 60/457,031
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: 60/457,730
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: 60/459,931
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/460,357
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/461,265
; PRIOR FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: 60/462,805
; PRIOR FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 60/468,139
; PRIOR FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: 60/464,886
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/471,200
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 2476
; SOFTWARE: PatentIn ver. 3.2
; SEQ ID NO 11
; LENGTH: 1213
; TYPE: DNA
; ORGANISM: Human severe acute respiratory system virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1213)
US-10-808-187-11
```

```
Query Match      100.0%; Score 25; DB 19; Length 1213;
Best Local Similarity 100.0%; Pred. No. 0.14; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;
```

```
QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 569 CAGAACGCTGTAGCTTCAAAAATCT 593
```

## RESULT 3

```
US-10-699-936-9
; Sequence 9, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 5262
; TYPE: DNA
; ORGANISM: SARS coronavirus Shanghai QXC
US-10-699-936-9
```

```
Query Match      100.0%; Score 25; DB 19; Length 5262;
Best Local Similarity 100.0%; Pred. No. 0.18; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;
```

```
QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 1984 CAGAACGCTGTAGCTTCAAAAATCT 2008
```

## RESULT 4

```
US-10-889-447-5
; Sequence 5, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ03
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-5
```

```
Query Match      100.0%; Score 25; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 0.24; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;
```

```
QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
|||||
Db 17445 CAGAACGCTGTAGCTTCAAAAATCT 17469
```

## RESULT 5

```
US-10-889-447-6
; Sequence 6, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
```

;; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
;; FILE REFERENCE: RTS-0685US  
;; CURRENT APPLICATION NUMBER: US/10/889,447  
;; PRIOR FILING DATE: 2004-07-12  
;; PRIOR APPLICATION NUMBER: 60/486,670  
;; PRIOR FILING DATE: 2003-07-12  
;; NUMBER OF SEQ ID NOS: 241  
;; SOFTWARE: FastSEQ for Windows Version 4.0  
;; SEQ ID NO 6  
;; LENGTH: 28920  
;; TYPE: DNA  
;; ORGANISM: SARS coronavirus isolate BJ04  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION:  
;; OTHER INFORMATION: n is any nucleotide  
US-10-889-447-6

Query Match 100.0%; Score 25; DB 19; Length 28920;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAACGCTGTAGCTTCAAAAATCT 25  
|||||  
Db 17445 CAGAACGCTGTAGCTTCAAAAATCT 17459

## RESULT 6

US-10-889-447-4  
;; Sequence 4, Application US/10889447  
;; Publication No. US20050075307A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Bennett, C. Frank  
;; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
;; FILE REFERENCE: RTS-0685US  
;; CURRENT APPLICATION NUMBER: US/10/889,447  
;; CURRENT FILING DATE: 2004-07-12  
;; PRIOR APPLICATION NUMBER: 60/486,670  
;; PRIOR FILING DATE: 2003-07-12  
;; NUMBER OF SEQ ID NOS: 241  
;; SOFTWARE: FastSEQ for Windows Version 4.0  
;; SEQ ID NO 4  
;; LENGTH: 29291  
;; TYPE: DNA  
;; ORGANISM: SARS coronavirus isolate BJ02  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION:  
;; OTHER INFORMATION: n is any nucleotide  
US-10-889-447-4

Query Match 100.0%; Score 25; DB 19; Length 29291;  
Best Local Similarity 100.0%; Pred. No. 0.25;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAACGCTGTAGCTTCAAAAATCT 25  
|||||  
Db 17771 CAGAACGCTGTAGCTTCAAAAATCT 17795

## RESULT 7

US-10-889-447-7  
;; Sequence 7, Application US/10889447  
;; Publication No. US20050075307A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Bennett, C. Frank  
;; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
;; FILE REFERENCE: RTS-0685US  
;; CURRENT APPLICATION NUMBER: US/10/889,447  
;; CURRENT FILING DATE: 2004-07-12  
;; PRIOR APPLICATION NUMBER: 60/486,670

;; PRIOR FILING DATE: 2003-07-12  
;; NUMBER OF SEQ ID NOS: 241  
;; SOFTWARE: FastSEQ for Windows Version 4.0  
;; SEQ ID NO 7  
;; LENGTH: 29430  
;; TYPE: DNA  
;; ORGANISM: SARS coronavirus isolate GZ01  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION:  
;; OTHER INFORMATION: n is any nucleotide  
US-10-889-447-7

Query Match 100.0%; Score 25; DB 19; Length 29430;  
Best Local Similarity 100.0%; Pred. No. 0.25;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAACGCTGTAGCTTCAAAAATCT 25  
|||||  
Db 17430 CAGAACGCTGTAGCTTCAAAAATCT 17454

## RESULT 8

US-10-839-729-15  
;; Sequence 15, Application US/10839729  
;; Publication No. US20050002953A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Jens Herold  
;; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES  
;; TITLE OF INVENTION: AND METHODS OF USE  
;; FILE REFERENCE: BIOBANK.013A  
;; CURRENT APPLICATION NUMBER: US/10/839,729  
;; PRIOR FILING DATE: 2004-05-04  
;; PRIOR APPLICATION NUMBER: 60/468703  
;; PRIOR FILING DATE: 2003-05-06  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: FastSEQ for Windows Version 4.0  
;; SEQ ID NO 15  
;; LENGTH: 29727  
;; TYPE: DNA  
;; ORGANISM: SARS Coronavirus  
US-10-839-729-15

Query Match 100.0%; Score 25; DB 18; Length 29727;  
Best Local Similarity 100.0%; Pred. No. 0.25;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGAACGCTGTAGCTTCAAAAATCT 25  
|||||  
Db 17718 CAGAACGCTGTAGCTTCAAAAATCT 17742

## RESULT 9

US-10-827-757-1  
;; Sequence 1, Application US/10827757  
;; Publication No. US20050004071A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Comper, Wayne  
;; TITLE OF INVENTION: Charged Polysaccharides Resistant To Lysosomal Degradation During  
;; TITLE OF INVENTION: Kidney Filtration And Renal Passage And Their Use To Treat Or  
;; TITLE OF INVENTION: Prevent Infection By Coronaviruses  
;; FILE REFERENCE: 11213-007-999  
;; CURRENT APPLICATION NUMBER: US/10/827,757  
;; CURRENT FILING DATE: 2004-04-20  
;; PRIOR APPLICATION NUMBER: 60/464,294  
;; PRIOR FILING DATE: 2003-04-21  
;; NUMBER OF SEQ ID NOS: 1  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 1  
;; LENGTH: 29727  
;; TYPE: DNA  
;; ORGANISM: SARS-related coronavirus (Urbani strain)  
US-10-827-757-1

```
Query Match      100.0%; Score 25; DB 18; Length 29727;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
      |||||||
Db      17718 CAGAACGCTGTAGCTTCAAAAATCT 17742

RESULT 10
US-10-889-447-8
; Sequence 8, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; PRIOR FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS coronavirus Urbani
US-10-889-447-8

Query Match      100.0%; Score 25; DB 19; Length 29727;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
      |||||||
Db      17718 CAGAACGCTGTAGCTTCAAAAATCT 17742

RESULT 11
US-10-699-936-1
; Sequence 1, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS coronavirus Urbani
US-10-699-936-1

Query Match      100.0%; Score 25; DB 19; Length 29727;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
      |||||||
Db      17718 CAGAACGCTGTAGCTTCAAAAATCT 17742

RESULT 12
```

```
US-10-839-729-17
; Sequence 17, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
; FILE REFERENCE: BIOBANK.013A
; CURRENT APPLICATION NUMBER: US/10/839,729
; CURRENT FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: 60/468703
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 29736
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-839-729-17

Query Match      100.0%; Score 25; DB 18; Length 29736;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
      |||||||
Db      17703 CAGAACGCTGTAGCTTCAAAAATCT 17727

RESULT 13
US-10-889-447-9
; Sequence 9, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 29736
; TYPE: DNA
; ORGANISM: SARS coronavirus CUHK-W1
US-10-889-447-9

Query Match      100.0%; Score 25; DB 19; Length 29736;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CAGAACGCTGTAGCTTCAAAAATCT 25
      |||||||
Db      17703 CAGAACGCTGTAGCTTCAAAAATCT 17727

RESULT 14
US-10-699-936-3
; Sequence 3, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
```

; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 29736  
; TYPE: DNA  
; ORGANISM: SARS coronavirus CUHK-WI  
US-10-699-936-3

Query Match 100.0%; Score 25; DB 19; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 0.25;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCTGTAGCTTCAAAAATCT 25  
|||||  
Db 17703 CAGACGCTGTAGCTTCAAAAATCT 17727

## RESULT 15

US-10-839-729-16  
; Sequence 16, Application US/10839729  
; Publication No. US2005002953A1  
; GENERAL INFORMATION:  
; APPLICANT: Jens Herold  
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES  
; FILE REFERENCE: BIOBANK.013A  
; CURRENT APPLICATION NUMBER: US/10/839,729  
; CURRENT FILING DATE: 2004-05-04  
; PRIOR APPLICATION NUMBER: 60/468703  
; PRIOR FILING DATE: 2003-05-06  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 29742  
; TYPE: DNA  
; ORGANISM: SARS Coronavirus  
US-10-839-729-16

Query Match 100.0%; Score 25; DB 18; Length 29742;  
Best Local Similarity 100.0%; Pred. No. 0.25;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACGCTGTAGCTTCAAAAATCT 25  
|||||  
Db 17718 CAGACGCTGTAGCTTCAAAAATCT 17742

Search completed: May 16, 2005, 14:33:44  
Job time : 1187.16 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:01:20 ; Search time 2823.44 Seconds

(without alignments)  
337.038 Million cell updates/sec

Title: US-10-808-187A-2471

Perfect score: 25

Sequence: 1 cagaacgtgtagcttcaaaaatct 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20.2	80.8	826	8	BH503697 BOHMB53TR
C 2	19.2	76.8	759	7	CF347362 AGENCOURT
C 3	18.8	75.2	398	8	B34627 HS-1025-A1-
C 4	18.8	75.2	611	9	CC962406 B0IDS81TF
C 5	18.6	74.4	319	5	BQ533690 H4023F11-
C 6	18.6	74.4	345	9	CL239830 ZMMBB057
C 7	18.6	74.4	384	4	BM121685 L0500D10
C 8	18.6	74.4	414	1	AA771415 vm44e01.r
C 9	18.6	74.4	453	8	AQ723105 HS 2083.A
C 10	18.6	74.4	457	2	BB670372 B5670372
C 11	18.6	74.4	459	1	AA420054 VF50C02.r
C 12	18.6	74.4	522	1	AI503985 vm44e01.x
C 13	18.6	74.4	542	8	AQ663564 HS 5476.A
C 14	18.6	74.4	545	8	AQ080522 HS 5298.B
C 15	18.6	74.4	673	8	AZ734853 RPI-24-1
C 16	18.6	74.4	687	8	BH956751 odh04b09.
C 17	18.6	74.4	703	4	BJ704394 BJ704394
C 18	18.6	74.4	705	4	BJ721525 BJ721525
C 19	18.6	74.4	714	4	BJ715645 BJ715645
C 20	18.6	74.4	755	1	AI604651 vm44e01.Y
C 21	18.6	74.4	778	8	BH428629 BOHMV77TF
C 22	18.6	74.4	788	4	BJ718255 BJ718255
C 23	18.6	74.4	789	8	BH561107 BOCQ758TF
C 24	18.6	74.4	808	6	CA459517 AGENCOURT

25	18.6	74.4	829	6	CB591647	CB591647	AGENCOURT
26	18.6	74.4	888	7	CV291083	CV291083	af01-10m
C 27	18.6	74.4	1112	3	AK009747	AK009747	Mus muscu
C 28	18.6	74.4	3844	3	AK036660	AK036660	Mus muscu
C 29	18.4	73.6	467	6	CB098886	CB098886	ks14c05.Y
C 30	18.4	73.6	517	6	CB099218	CB099218	ks07g05.Y
C 31	18.4	73.6	777	7	CB206425	CB206425	Tor6855.G
C 32	18.2	72.8	228	9	CC798330	CC798330	SALK_1461
C 33	18.2	72.8	244	1	AA034344	AA034344	zk20f01.r
C 34	18.2	72.8	293	1	AA430864	AA430864	RRAMCA363
C 35	18.2	72.8	307	8	BZ381548	BZ381548	SALK_1168
C 36	18.2	72.8	313	8	BZ381547	BZ381547	SALK_1168
C 37	18.2	72.8	314	8	BZ381602	BZ381602	SALK_1169
C 38	18.2	72.8	333	1	AI662572	AI662572	ms97g03.x
C 39	18.2	72.8	382	6	CB116811	CB116811	K-EST0161
C 40	18.2	72.8	395	5	BY375756	BY375756	BY375756
C 41	18.2	72.8	403	7	CN574017	CN574017	rc47f04.x
C 42	18.2	72.8	429	9	CE774308	CE774308	tigr-gss-
C 43	18.2	72.8	461	5	BY451803	BY451803	BY451803
C 44	18.2	72.8	469	2	AW213684	AW213684	uo90c05.x
C 45	18.2	72.8	470	8	AQ278670	AQ278670	CITBI-E1-

## ALIGNMENTS

RESULT 1  
BH503697/c  
LOCUS  
DEFINITION  
BOHMB53TR BOHM Brassica oleracea genomic clone BOHMB53, genomic survey sequence.  
ACCESSION  
BH503697  
VERSION  
BH503697.1  
KEYWORDS  
GSS.  
SOURCE  
Brassica oleracea  
ORGANISM  
Brassica oleracea  
REFERENCE  
1 (bases 1 to 826)  
AUTHORS  
Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
TIGR  
TITLE  
Whole genome shotgun sequencing of Brassica oleracea  
JOURNAL  
Unpublished (2001)  
COMMENT  
Other GSSs: BOHMB53TF  
Contact: Chris Town

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1. 826

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="TO1000DH3"

/db\_xref="taxon:3712"

/clone="BOHMB53"

/clone\_lib="BOHM"

/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared

genomic DNA inserted into pHOS1 using BstXI linkers"

## ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 826;

Best Local Similarity 88.0%; Pred. No. 1.6e+02;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CAGAAGCGTGTAGCTTCAAAAATCT 25

DB 115 CAGAAGCGTGTAGCTTCAAAAACCT 91

```

RESULT 2
CF347362
LOCUS
DEFINITION AGENCOURT_15223823 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:7002298
5' mRNA sequence.
ACCESSION CF347362
VERSION CF347362.1 GI:33790161
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 759)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14696 row: d column: 08
High quality sequence stop: 679.
Location/Qualifiers
1..759
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7002298"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available
(NIH_ZGC_7). Library was constructed by Open Biosystems
(Huntsville, AL)."
```

ORIGIN

Query Match 76.8%; Score 19.2; DB 7; Length 759;  
Best Local Similarity 87.5%; Pred. No. 4.7e+02;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATC 24  
||||| ||||| ||||| ||||| |||||  
DB 733 CAGAACCTGGAGCTTCAGAAATC 756

RESULT 3  
B34627/c  
LOCUS  
DEFINITION HS-1025-A1-D08-MF abi CIT Human Genomic Sperm Library C Homo sapiens genomic clone Plate-CT 804 Col=15 Row=G, genomic survey sequence.  
ACCESSION B34627  
VERSION B34627.1 GI:2533996  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 398)
Mahairas.G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors
Unpublished (1997)
Contact: Mahairas GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: krackron@u.washington.edu
Sequence Tagged Connector
Plate: CT 804 row: G column: 15
Class: BAC ends
High quality sequence stop: 398.
Location/Qualifiers
1..398
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=CT 804 Col=15 Row=G"
/sex="M"
/clone_lib="CIT Human Genomic Sperm Library C"
/notes="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
E-Coli DH10B"
```

ORIGIN

Query Match 75.2%; Score 18.8; DB 8; Length 398;  
Best Local Similarity 90.9%; Pred. No. 6.8e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACGCTGTAGCTTCAAAAAT 23  
||||| ||||| ||||| ||||| |||||  
DB 153 AGATCACTGTAGCTTCAAAAAT 132

RESULT 4  
CC962406 611 bp DNA linear GSS 18-AUG-2003  
BOIDS81TF BO\_1.4\_1.6 KB nuc Brassica oleracea genomic clone  
BOIDS81, genomic survey sequence.  
ACCESSION CC962406  
VERSION CC962406.1 GI:33810762  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
Whole genome shotgun sequencing of Brassica oleracea  
Unpublished (2001)  
Other\_GSSs: BOIDS81TR  
Contact: Chris Town  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TP  
Class: sheared ends.  
Location/Qualifiers
1..611
/organism="Brassica oleracea"
/mol\_type="genomic DNA"
/strain="T0100DH3"
/db\_xref="taxon:3712"
/clone="BOIDS81"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source

```

/clone_lib="BO 1.4.1.6 KB nuc"
/notes="vector: PHOS2; Site_1: BstXI; 1.4-1.6 kb sheared
nuclear DNA inserted into PHOS2 using BstXI linkers"

ORIGIN
Query Match      75.2%; Score 18.8; DB 9; Length 611;
Best Local Similarity 90.9%; Pred. No. 7e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AGAAGCTGTAGCTTCAAAAAT 23
|||||  |||||||  |||||||  |||
Db 196 AGAAGCTGTAGCTTCAATAAT 217

RESULT 5
BQ553690
LOCUS
DEFINITION      BQ553690 319 bp mRNA linear EST 20-JUN-2002
H4023F11-3 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
ACCESSION      BQ553690
VERSION        BQ553690
KEYWORDS       EST.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 319)
VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G.,
Martin,P.R., Staggs,C.A., Bassey,U., Aiba,K., Hamatani,T.,
Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.
Assembly, verification, and initial annotation of NIA 7.4K mouse
cDNA clone set
Genome Res. 12 (12), 1993-2003 (2002)
22354164
12466305
Contact: Yong Qian
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html for details.
Plate: H4023 row: F column: 11
Seq primer: -21M13 Forward
High quality sequence stop: 319
POLYA=Yes.

FEATURES
source
1..319
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="niaEST:H4023F11-3"
/db_xref="taxon:10090"
/clone="H4023F11"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/clone_lib="NIA Mouse 7.4K cDNA Clone Set"
/notes="Vector: pSPORT1; Site_1: Sali; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."

ORIGIN
Query Match      74.4%; Score 18.6; DB 5; Length 319;
Best Local Similarity 84.0%; Pred. No. 8.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAGAAGCTGTAGCTTCAAAAATCT 25
|||||  |||||||  |||||||  |||
Db 231 CAGAACCCTTAGCTTCAAAAATAT 255

RESULT 6

```

```

CL239830
LOCUS
DEFINITION      CL239830 345 bp DNA linear GSS 15-JAN-2004
ZMMBB0576D01r ZMMBB (HindIII) zea may's genomic clone
ZMMBB0576D01 3', genomic survey sequence.
ACCESSION      CL239830
VERSION        CL239830
KEYWORDS       GSS.
SOURCE         Zea mays
ORGANISM       Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 345)
Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
Sequencing of the maize genome at PCR (2003c)
Unpublished (2003)
Contact: Bharti,A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 66.
Location/Qualifiers
1..345
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBB0576D01"
/lab_host="E. coli DH10B"
/clone_lib="ZMMBB (HindIII)"
/notes="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match      74.4%; Score 18.6; DB 9; Length 345;
Best Local Similarity 84.0%; Pred. No. 8.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAGAAGCTGTAGCTTCAAAAATCT 25
|||||  |||||||  |||||||  |||
Db 57 CATAACTCTAGCTTAAGAATCT 81

RESULT 7
BM121685
LOCUS
DEFINITION      BM121685 384 bp mRNA linear EST 09-JUN-2003
L0500D10-3 NIA Mouse Newborn Heart cDNA Library Mus musculus cDNA
clone L0500D10 3', mRNA sequence.
ACCESSION      BM121685
VERSION        BM121685
KEYWORDS       EST.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 384)
Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
21429098
11544199
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health

```

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsun.grc.nia.nih.gov

Plate: L0500 row: D column: 10

Seq primer: -21M13 Forward

High quality sequence stop: 384

POLYA=Yes

#### FEATURES

source location/Qualifiers

1. .384  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="niaEST:L0500D10-3"  
/db\_xref="taxon:10090"  
/clone="L0500D10"  
/tissue\_type="Newborn Heart"  
/dev\_stage="Newborn"  
/lab\_host="DH10B"  
/note="Organ: heart; Vector: pSPORT1 (Invitrogen); Site: 1: Sali; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). Double-stranded cDNAs were synthesized with an Oligo (dT) primer [Invitrogen: 5'-  
pGACTAGTCTAGATCGGAGCGCCCTTTTCTTTT-3'] from  
24.9 microgram of total RNA, treated with T4 DNA  
polymerase, and purified by ethanol-precipitation. The  
cDNAs were ligated to Lone-linker LL-Sal3 (Ref.  
Development 127:1737-1749 (2000) [PMID: 10725249]),  
purified by phenol/chloroform, and separated from free  
linkers by Centricon 100. Then, the cDNAs were digested  
with Sali and NotI enzymes, and cloned into Sali and NotI  
site of pSPORT1 plasmid vector. The DH10B E. coli host was  
transformed with ligation mixture by the chemical method.  
The average insert size is about 1.8 kb. The library was  
constructed by Yulan Piao (NIA)."

#### ORIGIN

Query Match 74.4%; Score 18.6; DB 4; Length 384;  
Best Local Similarity 84.0%; Pred. No. 8.4e+02;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25

Db 358 CAGAACCTCTAGCTCCAAAATAT 382

#### RESULT 8

AA771415/c

LOCUS

DEFINITION vm44e01.r1 Stratagene mouse diaphragm (#937303) Mus musculus cDNA  
clone IMAGE:1001112 5' similar to gb:L35933 Mouse erythrocyte  
protein 4.2 (MOUSE);, mRNA sequence.

ACCESSION AA771415

VERSION AA771415.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1996)

Contact: Marra M/Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:565328

Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 406.

#### FEATURES

source location/Qualifiers

1. .414  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1001112"  
/tissue\_type="diaphragm"  
/dev\_stage="adult"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Organ: diaphragm; Vector: pBluescript SK-; Site 1:  
EcoRI; Site 2: XhoI; Cloned unidirectionally from mRNA  
prepared from diaphragm muscle. Primer: Oligo dT. Average  
insert size: 1.5 kb. Uni-ZAP XR Vector; ~5' adaptor  
sequence: 5' GAATTCGCACGAG 3' ~3' adaptor sequence: 5'  
CTCAGATTTTTTTTTTTTTT 3'"

#### ORIGIN

Query Match 74.4%; Score 18.6; DB 1; Length 414;  
Best Local Similarity 84.0%; Pred. No. 8.4e+02;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25

Db 412 CAGAACCTCTAGCTCCAAAATAT 388

#### RESULT 9

LOCUS

DEFINITION HS\_2083\_A2\_D09\_MR CIT Approved Human Genomic Sperm Library D Homo  
sapiens genomic clone Plate=2083 Col=18 Row=G, genomic survey  
sequence.

ACCESSION AQ723105

VERSION AQ723105.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

10449764

10449764

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.htsc.washington.edu

Plate: 2083 row: G column: 18

Seq primer: M13 Reverse

Class: BAC ends

High quality sequence stop: 453.

Location/Qualifiers

1. .453

/organism="Homo sapiens"

/mol\_type="genomic DNA"

#### FEATURES

source

```

/db xref="taxon:9606"
/clone="plate=2083 Col=18 Row=G"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelcBAC11; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match      74.4%; Score 18.6; DB 8; Length 453;
Best Local Similarity 84.0%; Pred. No. 8.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
    ||||| ||||| ||||| ||||| |||||
Db 227 CAAAACACTTAACTTCAAAAATCT 251

RESULT 10
BB670372/c
LOCUS BB670372 457 bp mRNA linear EST 05-OCT-2001
DEFINITION BB670372 RIKEN full-length enriched, 17 days embryo head Mus
            musculus cDNA clone 3322402N08 3', mRNA sequence.
ACCESSION BB670372.1 GI:15969593
VERSION BB670372.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
        Akimura T., Arakawa T., Carninci P., Furuno M., Hanagaki T.,
        Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Imotani K.,
        Ishii Y., Ito M., Kawai J., Kojima Y., Konno H., Kouda M.,
        Matsuyama T., Nakamura M., Nishi K., Nomura K., Numasaki R.,
        Okazaki Y., Okido T., Saito R., Sakai C., Sakai K., Sakazume M.,
        Sasaki D., Sato K., Shibata K., Shinagawa A., Shiraki T.,
        Sogabe Y., Suzuki H., Tagawa A., Takahashi F., Takaku-Akaira S.,
        Tanaka T., Tomaru A., Toyota T., Watahiki A., Yasunishi A.,
        Muramatsu M. and Hayashizaki Y.
        RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura T., et al.
        2001)
        Unpublished (2001)
        Contact: Yoshihide Hayashizaki
        Laboratory for Genome Exploration Research Group, RIKEN Genomic
        Sciences Center (GSC), Yokohama Institute
        The Institute of Physical and Chemical Research (RIKEN)
        1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
        Tel: 81-45-503-9222
        Fax: 81-45-503-9216
        Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
        Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K.,
        Itoh M., Konno H., Okazaki Y., Muramatsu M. and Hayashizaki Y.
        Normalization and subtraction of cap-trapper-selected cDNAs to
        prepare full-length cDNA libraries for rapid discovery of new
        genes. Genome Res. 10 (10), 1617-1630 (2000)
        wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
        Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
        Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
        and Hayashizaki, Y.
        RIKEN integrated sequence analysis (RISA) system--384-format
        sequencing pipeline with 384 multicapillary sequencer. Genome Res.
        10 (11), 1757-1771 (2000)
        Konno H., Fukunishi Y., Shibata K., Itoh M., Carninci P.,
        Sugahara Y. and Hayashizaki Y.
        Computer-based methods for the mouse full-length cDNA
        encyclopedia: real-time sequence clustering for construction of a
        nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
        Please visit our web site (http://genome.gsc.riken.go.jp) for
        further details.
        e mouse tissues.
        Location/Qualifiers
        source
        1. .457
        /organism="Mus musculus"

ORIGIN

Query Match      74.4%; Score 18.6; DB 2; Length 457;
Best Local Similarity 84.0%; Pred. No. 8.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAAATCT 25
    ||||| ||||| ||||| ||||| |||||
Db 102 CAGAACCCCTCTAGCTCCAAAATAT 78

RESULT 11
AA420054
LOCUS AA420054 459 bp mRNA linear EST 16-OCT-1997
DEFINITION vF50c02.r1 Soares mouse NbMH Mus musculus cDNA clone IMAGE:847202
            5' similar to gb:L34808 Mus musculus (MOUSE);, mRNA sequence.
ACCESSION AA420054.1 GI:2081350
VERSION AA420054.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
        Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
        Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
        Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B.,
        Theising B., Wylie T., Lennon G., Soares B., Wilson R. and
        Waterston R.
        The WashU-HMI Mouse EST Project
        Unpublished (1996)
        Contact: Marra M/Mouse EST Project
        WashU-HMI Mouse EST Project
        Washington University School of Medicine
        4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
        Tel: 314 286 1800
        Fax: 314 286 1810
        Email: mouseest@watson.wustl.edu
        This clone is available royalty-free through LLNL; contact the
        IMAGE Consortium (info@image.llnl.gov) for further information.
        MGI:499354
        Seq primer: -28m13 rev2 ET from Amersham
        High quality sequence stop: 394.
        Location/Qualifiers
        source
        1. .459
        /organism="Mus musculus"
        /mol_type="mRNA"

```

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:847202"
/sex="male"
/tissue_type="heart"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mouse NMH"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAGCGGCGCGCGAGTATTTTATTTTATTTTATTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pRT3 vector. RNA
provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

ORIGIN
Query Match 74.4%; Score 18.6; DB 1; Length 459;
Best Local Similarity 84.0%; Pred. No. 8.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAATCT 25
||||| ||||| ||||| ||||| |||||
Db 366 CAGAACCTCTAGCTCCAAAATAT 390

RESULT 12
AL503985
LOCUS AI503985 522 bp mRNA linear EST 11-MAR-1999
DEFINITION vm4e01.x1 Stratagene mouse diaphragm (#937303) Mus musculus cDNA
clone IMAGE:100112 3' similar to gb:135933 Mouse erythrocyte
protein 4.2 (MOUSE); mRNA sequence.
ACCESSION AI503985
VERSION AI503985.1 GI:4401836
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:565328
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
High quality sequence stop: 448.
Location/Qualifiers
1..522
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:100112"
/tissue_type="diaphragm"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse diaphragm (#937303)"
/note="Organ: diaphragm; Vector: pBluescript SK-; Site_1:

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:847202"
/sex="male"
/tissue_type="heart"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mouse NMH"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAGCGGCGCGCGAGTATTTTATTTTATTTTATTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pRT3 vector. RNA
provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

ORIGIN
Query Match 74.4%; Score 18.6; DB 1; Length 459;
Best Local Similarity 84.0%; Pred. No. 8.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAATCT 25
||||| ||||| ||||| ||||| |||||
Db 366 CAGAACCTCTAGCTCCAAAATAT 390

RESULT 13
AQ663564/c
LOCUS AQ663564 542 bp DNA linear GSS 23-JUN-1999
DEFINITION HS_5476_A2_C09_T7A_RPCI-11 Human Male BAC library Homo sapiens
genomic_clone Plate=1052 Col=18 Row=E, genomic survey sequence.
ACCESSION AQ663564
VERSION AQ663564.1 GI:5171332
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 542)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htrsc.washington.edu
Plate: 1052 row: E column: 18
Seq primer: T7
Class: BAC ends
High quality sequence stop: 542.
Location/Qualifiers
1..542
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1052 Col=18 Row=E"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

FEATURES
source
1..542
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1052 Col=18 Row=E"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

ORIGIN
Query Match 74.4%; Score 18.6; DB 8; Length 542;
Best Local Similarity 84.0%; Pred. No. 8.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

EcoRI; Site 2: XhoI; Cloned unidirectionally from mRNA
prepared from diaphragm muscle. Primer: Oligo dT. Average
insert size: 1.5 kb. Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGACGAG 3' -3' adaptor sequence: 5'
CTCAGATTTTTTTTTTTTTTTT 3'

ORIGIN
Query Match 74.4%; Score 18.6; DB 1; Length 522;
Best Local Similarity 84.0%; Pred. No. 8.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAGAACGCTGTAGCTTCAAAATCT 25
||||| ||||| ||||| ||||| |||||
Db 359 CAGAACCTCTAGCTCCAAAATAT 383

RESULT 13
AQ663564/c
LOCUS AQ663564 542 bp DNA linear GSS 23-JUN-1999
DEFINITION HS_5476_A2_C09_T7A_RPCI-11 Human Male BAC library Homo sapiens
genomic_clone Plate=1052 Col=18 Row=E, genomic survey sequence.
ACCESSION AQ663564
VERSION AQ663564.1 GI:5171332
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 542)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htrsc.washington.edu
Plate: 1052 row: E column: 18
Seq primer: T7
Class: BAC ends
High quality sequence stop: 542.
Location/Qualifiers
1..542
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1052 Col=18 Row=E"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

FEATURES
source
1..542
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1052 Col=18 Row=E"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

ORIGIN
Query Match 74.4%; Score 18.6; DB 8; Length 542;
Best Local Similarity 84.0%; Pred. No. 8.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 21:16:45 ; Search time 597.656 Seconds  
(without alignments)  
2026.885 Million cell updates/sec

Title: US-10-808-187A-2472  
Perfect score: 25  
Sequence: 1 tcagaacctgtgatcaacacag 25

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hcg: \*  
3: gb\_in: \*  
4: gb\_cm: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	25	100.0	626	14	AY451928 SARS CORO
C 2	25	100.0	626	14	AY451929 SARS CORO
C 3	25	100.0	626	14	AY451930 SARS CORO
C 4	25	100.0	626	14	AY451931 SARS CORO
C 5	25	100.0	626	14	AY451932 SARS CORO
C 6	25	100.0	626	14	AY451933 SARS CORO
C 7	25	100.0	626	14	AY451934 SARS CORO
C 8	25	100.0	626	14	AY451935 SARS CORO
C 9	25	100.0	626	14	AY451936 SARS CORO
C 10	25	100.0	626	14	AY451937 SARS CORO
C 11	25	100.0	626	14	AY451938 SARS CORO
C 12	25	100.0	626	14	AY451939 SARS CORO
C 13	25	100.0	626	14	AY451940 SARS CORO
C 14	25	100.0	626	14	AY451941 SARS CORO
C 15	25	100.0	626	14	AY451942 SARS CORO
C 16	25	100.0	626	14	AY451943 SARS CORO
C 17	25	100.0	626	14	AY451944 SARS CORO
C 18	25	100.0	626	14	AY451945 SARS CORO
C 19	25	100.0	646	14	AY268070 SARS CORO

C 20	25	100.0	677	14	AY443086S05	AY443090 SARS CORO
C 21	25	100.0	6067	14	AY534762S3	AY534764 SARS CORO
C 22	25	100.0	8796	14	AY534758S2	AY534759 SARS CORO
C 23	25	100.0	13471	14	AY304490	AY304490 SARS CORO
C 24	25	100.0	13471	14	AY304492	AY304492 SARS CORO
C 25	25	100.0	26333	14	AY286320	AY286320 SARS CORO
C 26	25	100.0	29013	14	AY463060	AY463060 SARS CORO
C 27	25	100.0	29350	14	AY394999	AY394999 SARS CORO
C 28	25	100.0	29350	14	AY395000	AY395000 SARS CORO
C 29	25	100.0	29350	14	AY395001	AY395001 SARS CORO
C 30	25	100.0	29350	14	AY395002	AY395002 SARS CORO
C 31	25	100.0	29433	14	AY394977	AY394977 SARS CORO
C 32	25	100.0	29530	14	AY394985	AY394985 SARS CORO
C 33	25	100.0	29573	14	AY338174	AY338174 SARS CORO
C 34	25	100.0	29573	14	AY338175	AY338175 SARS CORO
C 35	25	100.0	29573	14	AY348314	AY348314 SARS CORO
C 36	25	100.0	29577	14	AY559094	AY559094 SARS CORO
C 37	25	100.0	29592	14	AY463059	AY463059 SARS CORO
C 38	25	100.0	29620	14	AY395004	AY395004 SARS CORO
C 39	25	100.0	29640	14	AY394978	AY394978 SARS CORO
C 40	25	100.0	29645	14	AY394979	AY394979 SARS CORO
C 41	25	100.0	29646	14	AY394982	AY394982 SARS CORO
C 42	25	100.0	29647	14	AY395003	AY395003 SARS CORO
C 43	25	100.0	29661	14	AY559086	AY559086 SARS CORO
C 44	25	100.0	29665	14	AY394988	AY394988 SARS CORO
C 45	25	100.0	29670	14	AY559082	AY559082 SARS CORO

ALIGNMENTS

RESULT 1  
AY451928/c  
LOCUS  
DEFINITION SARS coronavirus TW-HPI isolate TW-HPI\_SC18 replicase 1B gene,  
partial cds.  
ACCESSION AY451928  
VERSION AY451928.1 GI:42741328  
KEYWORDS  
SOURCE SARS coronavirus TW-HPI  
ORGANISM SARS coronavirus TW-HPI  
REFERENCE 1 (bases 1 to 626)  
AUTHORS Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus.  
TITILE Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J., Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.  
JOURNAL Molecular Epidemiology of SARS in Taiwan - Development of a Simple Method for Tracing the Origin and Dissemination of SARS  
REFERENCE 2 (bases 1 to 626)  
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J., Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.  
TITILE Direct Submission  
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS Prevention and Research Center, 155, Li-Nong St., Sec.2, Feitou, Taipei, Taiwan 112, R.O.C.  
FEATURES  
source Location/Qualifiers  
1. .626  
/organism="SARS coronavirus TW-HPI"  
/mol\_type="genomic RNA"  
/isolate="TW-HPI\_SC18"  
/db\_xref="taxon:264371"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AA544814.1"  
/db\_xref="GI:42741329"  
/translation="ISMATNYDLSSVVRNARLAKHYVYIGDPAQLPAPRTLLTKGTLEP  
EYFNSCRMLKMTIGDPMFLGTCRCPCAEIVDTVSALVYNKLKAHKDQAQCQFMFVK  
GVITHDVSSAINRPQIGVVREPLTRNPARKAVFISPYNSQNAVASKILGLPTQTVD  
SGSEYDVYVIFTTETATSCNVRNVAITRAKIGILCIMSDDRLYD"

ORIGIN

```

Query Match      100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAACAG 25
|||||
Db 494 TCAGAACCTGTGATGAATCAACAG 470

RESULT 2
AY451929/c
LOCUS SARS coronavirus TW-HP2 isolate TW-HP2_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451929
VERSION AY451929.1 GI:42741330
KEYWORDS
SOURCE
ORGANISM SARS coronavirus TW-HP2
SARS coronavirus TW-HP2
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
TITLE
JOURNAL
AUTHORS
REFERENCE 2 (bases 1 to 626)
Db 494 TCAGAACCTGTGATGAATCAACAG 470

FEATURES
source
1..626
/organism="SARS coronavirus TW-HP2"
/mol_type="genomic RNA"
/isolate="TW-HP2_SC18"
/db_xref="taxon:264372"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44815.1"
/db_xref="GI:42741331"
/translation="ISMATNYDLSVNNARLRAKHVYVIGDPAQLPAPRTLTKGTLEP
EYFNSVCLMTIGDPMFLGTCRCPAEIVDTVSALVYDNKLKAHKDKSAOCFKMFYK
GVITHDVSSAINRPGVGVRELTNPARKAVFISPYNSQNAVASKILGLPTQTVDS
SQGSEYDYVIFTQTETTAHSCNVNRFVAITRAKIGILCIMSRLDLYD"

ORIGIN
Query Match      100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAACAG 25
|||||
Db 494 TCAGAACCTGTGATGAATCAACAG 470

RESULT 4
AY451931/c
LOCUS SARS coronavirus TW-HP4 isolate TW-HP4_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451931
VERSION AY451931.1 GI:42741334
KEYWORDS
SOURCE
ORGANISM SARS coronavirus TW-HP4
SARS coronavirus TW-HP4
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.
Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
TITLE
JOURNAL
AUTHORS
REFERENCE 2 (bases 1 to 626)
Db 494 TCAGAACCTGTGATGAATCAACAG 470

FEATURES
source
1..626
/organism="SARS coronavirus TW-HP4"
/mol_type="genomic RNA"
/isolate="TW-HP4_SC18"
/db_xref="taxon:264374"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44817.1"

CDS

```

```
/db_xref="GI:42741335"
/translation="ISMATNYDLSVNNARLRKHYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCRMLTKIGPDMFLGTCRCPAEIVDTVSALVDNKLKAHKDKSAQCFFMYK
GVITHDVSSAINRPOIGVREFLTRNPAMRKAVFISPSQNAVASKILGLPTQTVDS
SQGSEYDYVIFTQTTEHSCNVNRFNVAITRAKIGILCIMSDDRLDYD"

ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
|||||
Db 494 TCAGAACCCCTGTGATGAATCAACAG 470

RESULT 5
AY451932/c
LOCUS SARS coronavirus TW-JC2 isolate TW-JC2_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451932
VERSION AY451932.1 GI:42741336
KEYWORDS
SOURCE SARS coronavirus TW-JC2
ORGANISM SARS coronavirus TW-JC2
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y., and Chen, Y.-M. Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y., and Chen, Y.-M. Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1..626
/organism="SARS coronavirus TW-JC2"
/mol_type="genomic RNA"
/isolate="TW-JC2_SC18"
/db_xref="taxon:264375"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44818.1"
/db_xref="GI:42741337"
/translation="ISMATNYDLSVNNARLRKHYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCRMLTKIGPDMFLGTCRCPAEIVDTVSALVDNKLKAHKDKSAQCFFMYK
GVITHDVSSAINRPOIGVREFLTRNPAMRKAVFISPSQNAVASKILGLPTQTVDS
SQGSEYDYVIFTQTTEHSCNVNRFNVAITRAKIGILCIMSDDRLDYD"

ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
|||||
Db 494 TCAGAACCCCTGTGATGAATCAACAG 470

RESULT 6
AY451933/c
LOCUS SARS coronavirus TW-KC1 isolate TW-KC1_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451933
```

```
AY451933.1 GI:42741338
.
SARS coronavirus TW-KC1
SARS coronavirus TW-KC1
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
1 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y., and Chen, Y.-M. Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y., and Chen, Y.-M. Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1..626
/organism="SARS coronavirus TW-KC1"
/mol_type="genomic RNA"
/isolate="TW-KC1_SC18"
/db_xref="taxon:264376"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44819.1"
/db_xref="GI:42741339"
/translation="ISMATNYDLSVNNARLRKHYVIGDPAQLPAPRTLLTKGTLEP
EYFNSVCRMLTKIGPDMFLGTCRCPAEIVDTVSALVDNKLKAHKDKSAQCFFMYK
GVITHDVSSAINRPOIGVREFLTRNPAMRKAVFISPSQNAVASKILGLPTQTVDS
SQGSEYDYVIFTQTTEHSCNVNRFNVAITRAKIGILCIMSDDRLDYD"

ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
|||||
Db 494 TCAGAACCCCTGTGATGAATCAACAG 470

RESULT 7
AY451934/c
LOCUS SARS coronavirus TW-KC3 isolate TW-KC3_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451934
VERSION AY451934.1 GI:42741340
KEYWORDS SARS coronavirus TW-KC3
SOURCE SARS coronavirus TW-KC3
ORGANISM SARS coronavirus TW-KC3
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y., and Chen, Y.-M. Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,
Yang, J.-Y., Chen, H.-Y., and Chen, Y.-M. Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1..626
/organism="SARS coronavirus TW-KC3"
```



TITLE Direct Submission  
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou, Taipei, Taiwan 112, R.O.C.

FEATURES  
source  
1. .626  
/organism="SARS coronavirus TW-GD3"  
/mol\_type="genomic RNA"  
/isolate="TW-GD3\_SC18"  
/db\_xref="taxon:264380"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AA544823.1"  
/db\_xref="GI:42741347"  
/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLLTKGTLEP  
EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLKAHKDKSAQCQKMFYK  
GVTHDVSAINRPOIGVREPLTRNPARKAVFISPYNSONAVASKILGLPTQTVDS  
SQGSEYDYVIFTQTTEHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN  
Query Match 100.0%; Score 25; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25  
Db 494 TCAGAACCCCTGTGATGAATCAACAG 470

RESULT 11  
AY451938/c  
LOCUS SARS coronavirus TW-GD4 isolate TW-GD4\_SC18 replicase 1B gene,  
partial cds.  
ACCESSION AY451938.1 GI:42741348  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
SARS coronavirus TW-GD4  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE  
AUTHORS 1 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Molecular Epidemiology of SARS in Taiwan - Development of a Simple Method for Tracing the Origin and Dissemination of SARS  
Unpublished  
JOURNAL  
REFERENCE 2 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Direct Submission  
TITLE Submitted (28-OCT-2003) National Yang-Ming University, AIDS Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou, Taipei, Taiwan 112, R.O.C.

FEATURES  
source  
1. .626  
/organism="SARS coronavirus TW-GD4"  
/mol\_type="genomic RNA"  
/isolate="TW-GD4\_SC18"  
/db\_xref="taxon:264381"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AA544824.1"  
/db\_xref="GI:42741349"  
/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLLTKGTLEP  
EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLKAHKDKSAQCQKMFYK  
GVTHDVSAINRPOIGVREPLTRNPARKAVFISPYNSONAVASKILGLPTQTVDS  
SQGSEYDYVIFTQTTEHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN  
Query Match 100.0%; Score 25; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25  
Db 494 TCAGAACCCCTGTGATGAATCAACAG 470

RESULT 12  
AY451939/c  
LOCUS SARS coronavirus TW-GD5 isolate TW-GD5\_SC18 replicase 1B gene,  
partial cds.  
ACCESSION AY451939.1 GI:42741350  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
SARS coronavirus TW-GD5  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE  
AUTHORS 1 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Molecular Epidemiology of SARS in Taiwan - Development of a Simple Method for Tracing the Origin and Dissemination of SARS  
Unpublished  
JOURNAL  
REFERENCE 2 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Direct Submission  
TITLE Submitted (28-OCT-2003) National Yang-Ming University, AIDS Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou, Taipei, Taiwan 112, R.O.C.

FEATURES  
source  
1. .626  
/organism="SARS coronavirus TW-GD5"  
/mol\_type="genomic RNA"  
/isolate="TW-GD5\_SC18"  
/db\_xref="taxon:264382"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AA544825.1"  
/db\_xref="GI:42741351"  
/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLLTKGTLEP  
EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLKAHKDKSAQCQKMFYK  
GVTHDVSAINRPOIGVREPLTRNPARKAVFISPYNSONAVASKILGLPTQTVDS  
SQGSEYDYVIFTQTTEHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN  
Query Match 100.0%; Score 25; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25  
Db 494 TCAGAACCCCTGTGATGAATCAACAG 470

RESULT 13  
AY451940/c  
LOCUS SARS coronavirus TW-YM1 isolate TW-YM1\_SC18 replicase 1B gene,  
partial cds.  
ACCESSION AY451940  
VERSION AY451940.1 GI:42741352  
KEYWORDS  
SOURCE  
ORGANISM  
SARS coronavirus TW-YM1  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE  
AUTHORS 1 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,

Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25  
Db 494 TCAGAACCCCTGTGATGAATCAACAG 470

RESULT 12  
AY451939/c  
LOCUS SARS coronavirus TW-GD5 isolate TW-GD5\_SC18 replicase 1B gene,  
partial cds.  
ACCESSION AY451939.1 GI:42741350  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
SARS coronavirus TW-GD5  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE  
AUTHORS 1 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Molecular Epidemiology of SARS in Taiwan - Development of a Simple Method for Tracing the Origin and Dissemination of SARS  
Unpublished  
JOURNAL  
REFERENCE 2 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Direct Submission  
TITLE Submitted (28-OCT-2003) National Yang-Ming University, AIDS Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou, Taipei, Taiwan 112, R.O.C.

FEATURES  
source  
1. .626  
/organism="SARS coronavirus TW-GD5"  
/mol\_type="genomic RNA"  
/isolate="TW-GD5\_SC18"  
/db\_xref="taxon:264382"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AA544825.1"  
/db\_xref="GI:42741351"  
/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLLTKGTLEP  
EYFNSVCRMLKTIQPMFLGTCRCPAEIVDTVSALVYDNKLKAHKDKSAQCQKMFYK  
GVTHDVSAINRPOIGVREPLTRNPARKAVFISPYNSONAVASKILGLPTQTVDS  
SQGSEYDYVIFTQTTEHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

ORIGIN  
Query Match 100.0%; Score 25; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25  
Db 494 TCAGAACCCCTGTGATGAATCAACAG 470

RESULT 13  
AY451940/c  
LOCUS SARS coronavirus TW-YM1 isolate TW-YM1\_SC18 replicase 1B gene,  
partial cds.  
ACCESSION AY451940  
VERSION AY451940.1 GI:42741352  
KEYWORDS  
SOURCE  
ORGANISM  
SARS coronavirus TW-YM1  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE  
AUTHORS 1 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,

Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Molecular Epidemiology of SARS in Taiwan - Development of a Simple  
Method for Tracing the Origin and Dissemination of SARS

# JOURNAL REFERENCE

2 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,  
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Direct Submission  
Submitted (28-OCT-2003) National Yang-Ming University, AIDS  
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,  
Taipei, Taiwan 112, R.O.C.

# FEATURES source

1. 626  
/organism="SARS coronavirus TW-YM1"  
/mol\_type="genomic RNA"  
/isolate="TW-YM1\_SC18"  
/db\_xref="taxon:264383"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AAS44826.1"  
/db\_xref="GI:42741353"

# CDS

/translation="ISMATNYDLSVNNARLAKHYVYIGDPAQLPAPRTLTKGTLEP  
EYFNSVCLMKITIGDPLGTCRCRCPAEIVDTVSALVDNKLKAHKDSACQCFKMFYK  
GVITHDVSSALNRPQIGVREFLTNPAMRKAVFTSPYNSQNAVASKILGLPTQTVD  
SQSGSEYDVIPTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

# ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCAGAACCTGTGATCAACAG 25  
|||||  
Db 494 TCAGAACCTGTGATCAACAG 470  
|||||

RESULT 14  
AY451941/c  
LOCUS  
DEFINITION SARS coronavirus TW-YM2 isolate TW-YM2\_SC18 replicase 1B gene,  
partial cds.  
ACCESSION AY451941  
VERSION AY451941.1 GI:42741354  
KEYWORDS  
SOURCE SARS coronavirus TW-YM2  
ORGANISM SARS coronavirus TW-YM2  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

# REFERENCE AUTHORS

1 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,  
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Molecular Epidemiology of SARS in Taiwan - Development of a Simple  
Method for Tracing the Origin and Dissemination of SARS

# JOURNAL REFERENCE

2 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,  
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Direct Submission  
Submitted (28-OCT-2003) National Yang-Ming University, AIDS  
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,  
Taipei, Taiwan 112, R.O.C.

# FEATURES source

1. 626  
/organism="SARS coronavirus TW-YM2"  
/mol\_type="genomic RNA"  
/isolate="TW-YM2\_SC18"  
/db\_xref="taxon:264384"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AAS44827.1"  
/db\_xref="GI:42741355"

# CDS

/translation="ISMATNYDLSVNNARLAKHYVYIGDPAQLPAPRTLTKGTLEP  
EYFNSVCLMKITIGDPLGTCRCRCPAEIVDTVSALVDNKLKAHKDSACQCFKMFYK  
GVITHDVSSALNRPQIGVREFLTNPAMRKAVFTSPYNSQNAVASKILGLPTQTVD  
SQSGSEYDVIPTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

# ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCAGAACCTGTGATCAACAG 25  
|||||  
Db 494 TCAGAACCTGTGATCAACAG 470  
|||||

RESULT 15  
AY451942/c  
LOCUS  
DEFINITION SARS coronavirus TW-YM3 isolate TW-YM3\_SC18 replicase 1B gene,  
partial cds.  
ACCESSION AY451942  
VERSION AY451942.1 GI:42741356  
KEYWORDS  
SOURCE SARS coronavirus TW-YM3  
ORGANISM SARS coronavirus TW-YM3  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

# REFERENCE AUTHORS

1 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,  
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Molecular Epidemiology of SARS in Taiwan - Development of a Simple  
Method for Tracing the Origin and Dissemination of SARS

# JOURNAL REFERENCE

2 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,  
Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Direct Submission  
Submitted (28-OCT-2003) National Yang-Ming University, AIDS  
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,  
Taipei, Taiwan 112, R.O.C.

# FEATURES source

1. 626  
/organism="SARS coronavirus TW-YM3"  
/mol\_type="genomic RNA"  
/isolate="TW-YM3\_SC18"  
/db\_xref="taxon:264385"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AAS44828.1"  
/db\_xref="GI:42741357"

# CDS

/translation="ISMATNYDLSVNNARLAKHYVYIGDPAQLPAPRTLTKGTLEP  
EYFNSVCLMKITIGDPLGTCRCRCPAEIVDTVSALVDNKLKAHKDSACQCFKMFYK  
GVITHDVSSALNRPQIGVREFLTNPAMRKAVFTSPYNSQNAVASKILGLPTQTVD  
SQSGSEYDVIPTQTETTAHSCNVNRFNVAITRAKIGILCIMSDDRLYD"

# ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCAGAACCTGTGATCAACAG 25  
|||||  
Db 494 TCAGAACCTGTGATCAACAG 470  
|||||

Search completed: May 16, 2005, 02:00:23  
Job time : 597.656 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 17:17:00 ; Search time 320.703 Seconds  
(without alignments)  
461.466 Million cell updates/sec

Title: US-10-808-187A-2472

Perfect score: 25

Sequence: 1 tcagaacctgtgatgaatcaacag 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25	100.0	29751	12	ADJ39000 SARS coro
C 2	18.8	75.2	2400	8	ACC57972 Canola pr
C 3	18.6	74.4	436	2	AAX51917 Human sec
C 4	18.6	74.4	1067	12	ADL12960 Human ste
C 5	18.6	74.4	1169	6	ABZ11767 Human pol
C 6	18.6	74.4	1169	12	ADM44285 Novel hum
C 7	18.6	74.4	2386	4	AAI66415 Cytochrom
C 8	18.6	74.4	2950	4	AAH18169 Human cDN
C 9	18.6	74.4	4511	4	AAI04982 Human rep
C 10	18.6	74.4	4511	4	ABL97875 Human tes
C 11	18.6	74.4	5520	12	ADL12627 Human ste
C 12	18.6	74.4	135462	12	ADQ97061 Mouse can
C 13	18.6	74.4	192992	13	ABD32866 Mouse can
C 14	18.2	72.8	276	9	ADA60091 Soybean p
C 15	18.2	72.8	2716	12	ADQ63113 Novel hum
C 16	18.2	72.8	3755	4	ABL10352 Drosophil
C 17	18.2	72.8	3986	4	ABL10338 Drosophil
C 18	18	72.0	394	4	AAI82162 Human pol
C 19	18	72.0	1653	6	AAS94803 Human DNA
C 20	17.8	71.2	1689	10	ACC61068 Gene sequ

C 21	17.8	71.2	1689	10	ADK62847	Adk62847 Disease t
C 22	17.8	71.2	2393	12	ADK67447	Adk67447 Baker's y
C 23	17.6	70.4	193	5	ADL43131	Adl43131 Human cDN
C 24	17.6	70.4	348	5	AAS34286	Aas34286 Human cDN
C 25	17.6	70.4	466	4	ABA58142	Aba58142 Human foe
C 26	17.6	70.4	466	4	ABA58678	Aba58678 Human foe
C 27	17.6	70.4	466	4	AAI38357	Aai38357 Probe #70
C 28	17.6	70.4	466	4	AAI37749	Aai37749 Probe #64
C 29	17.6	70.4	466	4	AAK32530	Aak32530 Human bon
C 30	17.6	70.4	466	4	AAK31881	Aak31881 Human bon
C 31	17.6	70.4	466	4	AAK06811	Aak06811 Human bra
C 32	17.6	70.4	466	4	AAK06222	Aak06222 Human bra
C 33	17.6	70.4	466	4	ABS32241	Abs32241 Human liv
C 34	17.6	70.4	466	4	ABS31570	Abs31570 Human liv
C 35	17.6	70.4	466	6	ABS06642	Abs06642 Human gen
C 36	17.6	70.4	466	6	ABS07319	Abs07319 Human gen
C 37	17.6	70.4	591	5	ADL36747	Adl36747 Human ova
C 38	17.6	70.4	591	5	ADI71591	Adi71591 Human ova
C 39	17.6	70.4	1076	12	ADP04744	Adp04744 Sea squir
C 40	17.6	70.4	1729	2	AAZ28369	Aaz28369 Protein p
C 41	17.6	70.4	2418	6	ABN68560	Abn68560 Streptoco
C 42	17.6	70.4	2400	3	AAA88551	Aaa88551 Human dua
C 43	17.6	70.4	110000	6	ABN71527	Abn71527 16
C 44	17.6	70.4	110000	6	ABQ74964	Abq74964 4
C 45	17.6	70.4	110000	6	ABQ74964	Abq74964 5

ALIGNMENTS

RESULT 1

ADJ39000/c

ID ADJ39000 standard; DNA; 29751 BP.

XX

AC ADJ39000;

XX

DT 06-MAY-2004 (first entry)

XX

DE SARS coronavirus nucleotide sequence.

XX

KW small interfering RNA; siRNA; modified ribonucleotide;

KW viral replication inhibition; hepatitis C virus; HCV; hepatitis C;

KW antiinflammatory; hepatotropic; virucide; hepatitis A virus;

KW hepatitis D virus; hepatitis E virus; Ebola virus; influenza virus;

KW rotavirus; reovirus; retrovirus; poliovirus; human papilloma virus;

KW metapneumoniavirus; coronavirus; viral infection; gene; ds.

XX

OS SARS coronavirus.

XX

PN WO2004011647-A1.

XX

PD 05-FEB-2004.

XX

PF 25-JUL-2003; 2003WO-US023104.

XX

PR 26-JUL-2002; 2002US-0398605P.

XX

PA (CHIR ) CHIRON CORP.

XX

PI Han J, Seo MY, Houghton M;

XX

DR WPI; 2004-143862/14.

XX

PT New RNase resistant small interfering RNA, useful for treating viral

PT infections, e.g., hepatitis C, influenza virus or coronavirus infection.

XX

PS Example 10; Fig 3; 74pp; English.

XX

CC The present invention describes a small interfering RNA (siRNA) which

CC comprises a modified ribonucleotide, where the siRNA is resistant to

CC RNase and retains the ability to inhibit viral replication. Also

CC described: (1) inactivating a virus in a patient; (2) making a modified

CC siRNA that targets a nucleic acid sequence, in a virus; (3) a double-

CC stranded RNA molecule of 10-30 nucleotides that inhibits replication of  
 CC hepatitis C virus (HCV); (4) inducing targeted RNA interference toward  
 CC HCV in hepatic cells; (5) inhibiting replication of HCV; (6) a vector  
 CC comprising a DNA segment encoding the RNA molecule; (7) a host cell  
 CC comprising the vector of (6); (8) inhibiting replication of HCV in cells  
 CC carrying HCV; (9) treating hepatitis C in a subject; (10) a modified  
 CC siRNA molecule comprising a double-stranded RNA molecule of 10-30  
 CC nucleotides in length, which mediates RNA interference toward a target  
 CC agent or virus and is linked to at least one receptor-binding ligand; and  
 CC (11) inducing targeted RNA interference in a patient. The modified siRNA  
 CC molecules have antiinflammatory, hepatotropic and virucide activities.  
 CC The modified RNA molecules are useful for inactivating virus in mammalian  
 CC cells. The siRNAs are useful for treating hepatitis C virus, hepatitis A  
 CC virus, hepatitis D virus, hepatitis E virus, Ebola virus, influenza  
 CC virus, rotavirus, reovirus, retrovirus, poliovirus, human papilloma  
 CC virus, metapneumoniavirus or coronavirus infections. The methods of the  
 CC invention can be used to correct or compensate for cellular physiological  
 CC abnormalities involved in conferring susceptibility to viral infections  
 CC in patients and/or alleviate symptoms of a viral infection in patients.  
 CC The present sequence represents the SARS coronavirus nucleotide sequence,  
 CC which is used in an example from the present invention.

XX Sequence 29751 BP; 8478 A; 5941 C; 6188 G; 9144 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 12; Length 29751;  
 Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0

QY 1 TCAGAACCTGTGATGAATCAACAG 25

Db 17785 TCAGAACCTGTGATGAATCAACAG 17761

# RESULT 2

ID ACC57972 standard; cDNA; 2400 BP.

XX ACC57972;

XX 11-AUG-2003 (first entry)

DE Canola protein kinase stress-related polypeptide BnPK-4 cDNA.

XX BnPK-4; protein kinase stress-related polypeptide; PKSRP; enzyme;  
 KW transgenic plant; plant; stress tolerance; drought tolerance;  
 KW salt tolerance; cold tolerance; canola; gene; ss.

XX Brassica napus.

XX Key Location/Qualifiers  
 FH 48..2198  
 CDS /\*tag= a  
 FT /product= "BnPK-4"

XX WO2003040171-A2.

XX 15-MAY-2003.

XX 12-NOV-2002; 2002WO-US036374.

XX 09-NOV-2001; 2001US-0346096P.

XX (BADI ) BASF PLANT SCI GMBH.

XX Van Thiel N, Da Costa E SilvaO, Chen R;

XX WPI; 2003-441522/41.

XX P-PSDB; ABR42365.

XX New protein kinase stress-related polypeptide coding nucleic acid, useful  
 PT for producing transgenic plants with an increased tolerance to an  
 PT environmental stress, e.g. high salinity, as compared to a wild type  
 PT variety of the plant.

XX Claim 1; Page 92-93; 111pp; English.

XX The present sequence is a full-length cDNA encoding BnPK-4, a novel  
 CC canola protein kinase stress-related polypeptide (PKSRP). The cDNA was  
 CC identified on the basis of homology to Physcomitrella patens PK-3 (see  
 CC ACC57962) and PK-10 (see ACC57966) sequence. BnPK-4 is an example of  
 CC PKSRPs of the invention that are important for modulating a plant's  
 CC response to environmental stress. Over-expression of PKSRP coding nucleic  
 CC acids in a plant results in the plant's increased tolerance to  
 CC environmental stress. Transgenic monocot and dicot plants are provided  
 CC that show increased tolerance to high salinity, drought and low  
 CC temperature

XX Sequence 2400 BP; 702 A; 511 C; 500 G; 687 T; 0 U; 0 Other;

Query Match 75.2%; Score 18.8; DB 8; Length 2400;

Best Local Similarity 90.9%; Pred. No. 1e+02;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATGAATCAAC 23

Db 2166 CAGAACCTGTGATGAATCAAC 2145

# RESULT 3

AAAX51917/c

ID AAAX51917 standard; DNA; 436 BP.

XX AAAX51917;

XX 22-JUN-1999 (first entry)

DE Human secreted protein 5' EST SEQ ID NO: 131.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;

KW forensic; gene therapy; chromosome mapping; signal peptide;

KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;

KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;

XX thrombolytic; anti-inflammatory; tumour inhibition; ds.

XX Homo sapiens.

XX WO9906552-A2.

XX 11-FEB-1999.

XX 31-JUL-1998; 98WO-IB001236.

XX 01-AUG-1997; 97US-00905223.

XX (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Lacroix B;

XX WPI; 1999-153782/13.

XX P-PSDB; AAY13117.

XX New isolated brain-derived nucleic acids - used to develop products which  
 PT may have cytokine, immune, regulatory, haematopoiesis regulating, anti-  
 PT inflammatory or tumour inhibition activity.

XX Claim 1; Page 272; 577pp; English.

XX AAAX51787 to AAAX52019 represent 5' expressed sequence tags (ESTs) for  
 CC human secreted proteins, and encode the proteins given in AAY12987 to  
 CC AAY13219, respectively. The proteins given represent the signal peptide  
 CC and an N-terminal fragment of a secreted protein. The nucleic acid  
 CC sequences can be used for producing secreted human gene products. They  
 CC can also be used to develop products for diagnosis and therapy. The  
 CC proteins obtained may have cytokine activity, cell  
 CC proliferation/differentiation activity, haematopoiesis regulating





CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 1169 BP; 361 A; 198 C; 254 G; 356 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 6; Length 1169;  
Best Local Similarity 84.0%; Pred. No. 1.1e+02;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAACAG 25

Db 213 TCAGAACCATGGGATGTATCAGCAG 189

## RESULT 6

ADM44285/c

ID ADM44285 standard; cDNA; 1169 BP.

XX AC

XX ADM44285;

DT 03-JUN-2004 (first entry)

DE Novel human arginine-rich protein cDNA #649.

KW ss; gene; human; arginine-rich protein; cancer; inflammation;  
KW genetic disorder.

XX OS Homo sapiens.

XX KW (XUEA/) XUE A.

XX PA (DRMA/) DRMANAC R T.

XX XX Tang YT, Xue A, Drmanac RT;

XX WPI; 2004-238579/22.

XX PT New isolated arginine-rich protein-like polynucleotides and polypeptides,  
XX useful for diagnosing and/or treating conditions associated with aberrant  
XX activity of the arginine-rich polypeptides, such as cancer and  
XX inflammation.

XX PS Disclosure; SEQ ID NO 649; 51bp; English.

XX CC The invention relates to an isolated polynucleotide. The methods and  
XX compositions of the present invention are useful for the diagnosis and/or  
XX treatment of diseases or conditions associated with aberrant expression  
XX or activity of the arginine-rich protein-like polypeptides, such as  
XX cancer and inflammation. They can also be used in forensics, gene  
XX mapping, identification of mutations responsible for genetic disorders,  
XX CC and in assessing biodiversity. The present sequence represents a novel  
XX human arginine-rich protein cDNA.

XX SQ Sequence 1169 BP; 361 A; 198 C; 254 G; 356 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 12; Length 1169;

Best Local Similarity 84.0%; Pred. No. 1.1e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAACAG 25

Db 213 TCAGAACCATGGGATGTATCAGCAG 189

RESULT 7  
AAI66415/c  
ID AAI66415 standard; cDNA; 2386 BP.

XX AC AAI66415;

XX DT 04-DEC-2001 (first entry)

XX DE Cytochrome C oxidase COII protein 9 coding sequence.

XX KW Cytochrome C oxidase COII protein 9; cancer; HIV infection; gene therapy;  
XX ss.

XX OS Unidentified.

XX PN CN1300819-A.

XX PD 27-JUN-2001.

XX PF 21-DEC-1999; 99CN-00124290.

XX PR 21-DEC-1999; 99CN-00124290.

XX PA (UYFU-) UNIV FUDAN.

XX PI Mao Y, Xie Y;

XX DR WPI; 2001-530467/59.

XX DR P-PSDB; AAG78675.

XX PT Polypeptide-cytochrome C oxidase COII protein 9 and polynucleotide for  
XX coding this polypeptide.

XX PS Claim 6; Page 25-26 (Disclosure); 33pp; Chinese.

XX CC The present invention provides the protein and coding sequences of  
XX cytochrome C oxidase COII rptoien 9. The sequences can be used in the  
XX treatment of cancer and HIV infection. The present sequence is the coding  
XX sequence of the invention

XX SQ Sequence 2386 BP; 700 A; 402 C; 509 G; 775 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 4; Length 2386;

Best Local Similarity 84.0%; Pred. No. 1.3e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAACAG 25

Db 264 TCAGAACCATGGGATGTATCAGCAG 240

## RESULT 8

AAH18169/c

ID AAH18169 standard; cDNA; 2950 BP.

XX AC AAH18169;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:18065.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.





PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 08-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231124P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 25-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 27-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483232/52.  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
PT for preventing, diagnosing and/or treating testicular cancer.  
XX  
XX Disclosure; SEQ ID NO 2527; 766pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention  
XX  
SQ Sequence 4511 BP; 1177 A; 1009 C; 1055 G; 1270 T; 0 U; 0 Other;  
Query Match 74.4%; Score 18.6; DB 4; Length 4511;  
Best Local Similarity 84.0%; Pred. No. 1.4e+02;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 TCAGAACCCCTGTGATGAATCAACAG 25  
|||||  
Db 716 TCAGAACCCCTGTGATGAATGAGCAG 740  
RESULT: 11  
ADL12627/c  
ID ADL12627 standard; cDNA; 5520 BP.  
XX  
XX ADL12627;  
XX  
XX 06-MAY-2004 (first entry)  
XX



PT for diagnosing, preventing or treating cancers, especially lymphoma and  
 PT leukemia, or in screening for agents that modulate cancer.

XX disclosure; seqid 545; 310pp; English.

XX The invention relates to an isolated nucleic acid comprising at least 10  
 CC contiguous nucleotides of any of the 233 polynucleotide sequences given  
 CC in the specification, or its complement. The nucleic acids encode cancer-  
 CC associated proteins. Also included are an expression vector comprising  
 CC the isolated nucleic acid cited above, a host cell comprising the above  
 CC recombinant nucleic acid or expression vector, a microarray for detecting  
 CC a cancer-associated (CA) nucleic acid comprising at least one probe  
 CC comprising at least 10 contiguous nucleotides of any of the above-  
 CC mentioned nucleotide sequences, an isolated polypeptide (encoded within  
 CC an open reading frame of a CA sequence selected from any of the 95  
 CC polynucleotide sequences as mentioned in the specification, or its  
 CC complement), an isolated antibody, (or its antigen binding fragment) that  
 CC binds to the above polypeptide, a hybridoma that produces the above  
 CC monoclonal antibody, a pharmaceutical composition comprising the above  
 CC antibody and a pharmaceutical excipient, a kit for detecting cancer  
 CC cells (comprising the antibody cited above, methods for diagnosing cancer  
 CC or for detecting the presence or absence of cancer cells in an  
 CC individual, a method for inhibiting growth of cancer cells in an  
 CC individual, a method for delivering a therapeutic agent to cancer cells  
 CC in an individual, an electronic library comprising the above  
 CC polynucleotide or polypeptide (or their fragments), methods of screening  
 CC for anticancer activity or for a bioactive agent capable of modulating  
 CC the activity of a CA protein (CAP), methods for detecting cancer  
 CC associated with expression of a polypeptide in a test cell sample, a  
 CC method for treating cancers and a method for inhibiting the expression of  
 CC CA gene in a cell. The composition and methods are useful for detecting,  
 CC diagnosing, preventing and treating cancers, especially lymphoma and  
 CC leukemia. These may also be used in screening for agents that modulate  
 CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 192992 BP; 45285 A; 44727 C; 46554 G; 50925 T; 0 U; 5501 Other;

Query Match 74.4%; Score 18.6; DB 13; Length 192992;  
 Best Local Similarity 84.0%; Pred. No. 2.6e+02;  
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAGACCCCTGTGATGAATCAACAG 25

Db 93689 TCAGACCACTGTGATGAATCAACAG 93665

RESULT 14

ADA60091

ID ADA60091 standard; cDNA; 276 BP.

XX ADA60091;

XX 20-NOV-2003 (first entry)

XX Soybean pyrophosphatase. Fructose-6-phosphate phosphotransferase EST #4.

XX Plant; ss; EST; expressed sequence tag; transgenic; sucrose pathway;  
 KW triose phosphate isomerase; fructose 1; 6-bisphosphate aldolase;  
 KW fructose 1; 6-bisphosphate; fructose 6-phosphate 2-kinase;  
 KW phosphoglucose isomerase; vacuolar hydrogen translocating-pyrophosphatase;  
 KW pyrophosphate-dependent fructose-6-phosphate phosphotransferase;  
 KW invertase; sucrose synthase; hexokinase; fructokinase;  
 KW nucleoside diphosphate kinase-kinase; NDP;  
 KW glucose-6-phosphate 1-dehydrogenase; phosphoglucose mutase; UDP;  
 KW uridine diphosphate-glucose pyrophosphorylase; soybean.

XX Glycine max.

XX US2003135870-A1.

XX

PD 17-JUL-2003.  
 XX PF 26-JAN-1999; 99US-00237183.  
 XX XX  
 PR 24-NOV-1997; 97US-0067000P.  
 PR 09-DEC-1997; 97US-0069472P.  
 PR 27-JAN-1998; 98US-0072888P.  
 PR 10-FEB-1998; 98US-0074201P.  
 PR 10-FEB-1998; 98US-0074280P.  
 PR 10-FEB-1998; 98US-0074281P.  
 PR 10-FEB-1998; 98US-0074282P.  
 PR 12-FEB-1998; 98US-0074565P.  
 PR 12-FEB-1998; 98US-0074566P.  
 PR 12-FEB-1998; 98US-0074567P.  
 PR 19-FEB-1998; 98US-0074789P.  
 PR 19-FEB-1998; 98US-0075459P.  
 PR 19-FEB-1998; 98US-0075460P.  
 PR 19-FEB-1998; 98US-0075461P.  
 PR 19-FEB-1998; 98US-0075462P.  
 PR 19-FEB-1998; 98US-0075463P.  
 PR 19-FEB-1998; 98US-0075464P.  
 PR 06-MAR-1998; 98US-0076912P.  
 PR 09-MAR-1998; 98US-0077229P.  
 PR 09-MAR-1998; 98US-0077230P.  
 PR 09-MAR-1998; 98US-0077231P.  
 PR 18-MAR-1998; 98US-0078368P.  
 PR 07-APR-1998; 98US-0080844P.  
 PR 27-APR-1998; 98US-0083067P.  
 PR 29-APR-1998; 98US-0083386P.  
 PR 29-APR-1998; 98US-0083387P.  
 PR 29-APR-1998; 98US-0083388P.  
 PR 29-APR-1998; 98US-0083389P.  
 PR 29-APR-1998; 98US-0083390P.  
 PR 13-MAY-1998; 98US-0085222P.  
 PR 13-MAY-1998; 98US-0085223P.  
 PR 13-MAY-1998; 98US-0085224P.  
 PR 21-MAY-1998; 98US-0086183P.  
 PR 21-MAY-1998; 98US-0086184P.  
 PR 21-MAY-1998; 98US-0086185P.  
 PR 21-MAY-1998; 98US-0086186P.  
 PR 21-MAY-1998; 98US-0086187P.  
 PR 21-MAY-1998; 98US-0086188P.  
 PR 01-JUN-1998; 98US-0087422P.  
 PR 16-JUN-1998; 98US-0089524P.  
 PR 18-JUN-1998; 98US-0089793P.  
 PR 18-JUN-1998; 98US-0089810P.  
 PR 18-JUN-1998; 98US-0089814P.  
 PR 22-JUN-1998; 98US-0090170P.  
 PR 26-JUN-1998; 98US-0090928P.  
 PR 29-JUN-1998; 98US-0091035P.  
 PR 30-JUN-1998; 98US-0091405P.  
 PR 08-JUL-1998; 98US-0092036P.  
 PR 09-SEP-1998; 98US-0096667P.  
 PR 09-SEP-1998; 98US-0096670P.  
 PR 09-SEP-1998; 98US-0096697P.  
 PR 16-SEP-1998; 98US-0100672P.  
 PR 16-SEP-1998; 98US-0100673P.  
 PR 16-SEP-1998; 98US-0100674P.  
 PR 21-SEP-1998; 98US-0101130P.  
 PR 21-SEP-1998; 98US-0101131P.  
 PR 21-SEP-1998; 98US-0101132P.  
 PR 22-SEP-1998; 98US-0101343P.  
 PR 22-SEP-1998; 98US-0101344P.  
 PR 22-SEP-1998; 98US-0101347P.  
 PR 25-SEP-1998; 98US-0101508P.  
 PR 25-SEP-1998; 98US-0101707P.  
 PR 13-OCT-1998; 98US-0104123P.  
 PR 13-OCT-1998; 98US-0104124P.  
 PR 13-OCT-1998; 98US-0104126P.  
 PR 13-OCT-1998; 98US-0104127P.  
 PR 13-OCT-1998; 98US-0104128P.  
 PR 18-NOV-1998; 98US-0108996P.  
 PR 19-NOV-1998; 98US-0109018P.

```

PR 24-NOV-1998; 98US-00199129.
PR 08-DEC-1998; 98US-00210297.
PR 11-DEC-1998; 98US-0111981P.
PR 22-DEC-1998; 98US-0113224P.
PR 12-JAN-1999; 99US-00229413.
XX
PA (CHEI//) CHEIKH N.
PA (FISH//) FISHER D K.
PA (LIU//) LIU J.
XX
PI Cheikh N, Fisher DK, Liu J;
XX
XX WPI; 2003-688722/65.
XX
PT New maize or soybean enzymes and nucleic acid molecules associated with
PT the sucrose pathway, useful for genome mapping, gene identification and
PT analysis, plant breeding, or preparation of constructs for plant gene
PT expression.
XX
XX Claim 2; Page; 117pp; English.
XX
CC The invention relates to a substantially purified nucleic acid molecule
CC (appearing as ADA57847 - ADA60660 that encodes a maize or soybean enzyme
CC or its fragment, associated with the sucrose pathway selected from:
CC triose phosphate isomerase, fructose 1,6-bisphosphate aldolase, fructose
CC 1,6-bisphosphate, fructose 6-phosphate 2-kinase, phosphoglucose isomerase,
CC vacuolar hydrogen translocating-pyrophosphatase, pyrophosphate-dependent
CC fructose-6-phosphate phosphotransferase, invertase, sucrose synthase,
CC hexokinase, fructokinase, nucleoside diphosphate kinase (NDP)-kinase,
CC glucose-6-phosphate 1-dehydrogenase, phosphoglucomutase and uridine
CC diphosphate (UDP)-glucose pyrophosphorylase. Also included are a
CC substantially purified maize or soybean enzyme (or its fragment) and a
CC transformed plant having a nucleic acid molecule. Also disclosed as new
CC are purified antibodies capable of specifically binding to the maize or
CC soybean enzyme, determining a level or pattern of a plant sucrose pathway
CC enzyme in a plant cell or plant tissue, determining a mutation in a plant
CC whose presence is predictive of a mutation affecting the level or pattern
CC of a plant sucrose pathway enzyme, producing a plant containing an
CC overexpressed or reduced level of plant sucrose pathway enzyme, reducing
CC expression of a plant sucrose pathway enzyme in a plant and determining
CC an association between a polymorphism and a plant trait. The maize or
CC soybean enzymes and nucleic acid molecules are useful for genome mapping,
CC gene identification and analysis, plant breeding, or preparation of
CC constructs for plant gene expression and transgenic plants. The nucleic
CC acid molecules are also useful as markers or probes. The present sequence
CC is a soybean EST (expressed sequence tag) from a gene encoding a sucrose
CC pathway enzyme. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030135870.
XX
XX Sequence 276 BP; 83 A; 54 C; 71 G; 68 T; 0 U; 0 Other;
SQ
Query Match 72.8%; Score 18.2; DB 9; Length 276;
Best Local Similarity 87.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAAC 23
DB 8 TCAGAACCTGTGATGATCCAC 30
|||||
|||||

RESULT 15
ADQ63113/c
ID ADQ63113 standard; cDNA; 2716 BP.
XX
XX ADQ63113;
XX
XX 07-OCT-2004 (first entry)
XX
XX Novel human cDNA sequence #274.
XX
XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW

```

---

```

KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX
XX Homo sapiens.
XX
XX EP1440981-A2.
XX
XX 28-JUL-2004.
XX
XX 21-JAN-2004; 2004EP-00001196.
XX
XX 21-JAN-2003; 2003JP-00102206.
XX
XX 09-MAY-2003; 2003JP-00131392.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Nagai K, Irie R;
XX
XX WPI; 2004-535376/52.
XX
XX P-FSDB; ADQ63301.
XX
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX Claim 1; SEQ ID NO 274; 2449pp; English.
XX
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
XX polypeptides, sequences hybridizing to these nucleotides, sequences
XX encoding partial polypeptides and sequences having 70% or 90% identity to
XX the nucleotide and protein sequences. The nucleotides and polypeptides
XX are useful as diagnostic markers or therapeutic target for the diseases
XX or morbid states. They are also useful for treating osteoporosis,
XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,
XX dementia and various cancers. This sequence corresponds to a nucleotide
XX sequence of the invention.
XX
XX Sequence 2716 BP; 768 A; 637 C; 652 G; 659 T; 0 U; 0 Other;
SQ
Query Match 72.8%; Score 18.2; DB 12; Length 2716;
Best Local Similarity 87.0%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AGAACCTGTGATGAATCAACAG 25
DB 93 AGAACCTGTGCTGAATAACAG 71
|||||
|||||

Search completed: May 16, 2005, 01:09:21
Job time : 325.703 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:11:05 ; Search time 93.5547 Seconds  
(without alignments)  
437.251 Million cell updates/sec

Title: US-10-808-187A-2472

Perfect score: 25

Sequence: 1 tcgaaccctgtgatgaatcaacag 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*
  - 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
  - 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*
  - 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.6	74.4	436	3	US-08-905-223-168
C 2	18.6	74.4	1067	4	US-09-976-594-689
C 3	18.6	74.4	1169	4	US-09-799-451-649
C 4	18.6	74.4	5520	4	US-09-976-594-356
C 5	17.6	70.4	1729	3	US-09-045-973-6
C 6	17.6	70.4	45684	4	US-09-949-016-16539
C 7	17.6	70.4	786431	4	US-09-751-389-3
C 8	17.4	69.6	2234	4	US-09-949-016-3362
C 9	17.4	69.6	33908	4	US-09-949-016-15104
C 10	17.2	68.8	601	4	US-09-949-016-109231
C 11	17.2	68.8	57726	4	US-09-949-016-16430
C 12	17.2	68.8	154600	4	US-09-949-016-14757
C 13	17	68.0	6653	4	US-09-799-451-489
C 14	17	68.0	21706	3	US-08-961-527-36
C 15	17	68.0	524032	4	US-09-949-016-16928
C 16	17	68.0	524032	4	US-09-949-016-16929
C 17	17	68.0	524032	4	US-09-949-016-16930
C 18	17	68.0	524032	4	US-09-949-016-16931
C 19	17	68.0	529885	4	US-09-949-016-14340
C 20	17	68.0	529885	4	US-09-949-016-14341
C 21	17	68.0	529885	4	US-09-949-016-14342
C 22	17	68.0	529885	4	US-09-949-016-14343
C 23	17	68.0	529885	4	US-09-949-016-14344
C 24	17	68.0	529885	4	US-09-949-016-14345
C 25	17	68.0	529885	4	US-09-949-016-14346
C 26	17	68.0	529885	4	US-09-949-016-14347
C 27	17	68.0	640681	4	US-09-790-988-1

C 28	17	68.0	784019	4	US-09-949-016-14033
C 29	16.8	67.2	405	4	US-09-495-050A-213
C 30	16.8	67.2	601	4	US-09-949-016-76775
C 31	16.8	67.2	119801	4	US-09-949-016-13453
C 32	16.8	67.2	211049	4	US-09-949-016-15770
C 33	16.8	67.2	301828	4	US-09-949-016-13969
C 34	16.6	66.4	54	1	US-08-319-836B-1
C 35	16.6	66.4	54	5	PCT-US95-13142-1
C 36	16.6	66.4	63	1	US-08-319-836B-11
C 37	16.6	66.4	63	5	PCT-US95-13142-11
C 38	16.6	66.4	547	4	US-09-451-651-16
C 39	16.6	66.4	597	4	US-09-495-050A-72
C 40	16.6	66.4	601	4	US-09-949-016-37789
C 41	16.6	66.4	601	4	US-09-949-016-44271
C 42	16.6	66.4	601	4	US-09-949-016-155566
C 43	16.6	66.4	601	4	US-09-949-016-155567
C 44	16.6	66.4	601	4	US-09-949-016-155568
C 45	16.6	66.4	601	4	US-09-949-016-155569

ALIGNMENTS

RESULT 1

US-08-905-223-168/c  
; Sequence 168, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duerlet, Aymeric  
; APPLICANT: Lacroix, Bruno  
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
; NUMBER OF SEQUENCES: 503  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/905,223  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 168:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 436 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: DOUBLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: CDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Homo Sapiens  
; TISSUE TYPE: Brain  
; FEATURE:  
; NAME/KEY: sig peptide  
; LOCATION: 290..361  
; IDENTIFICATION METHOD: Von Heijne matrix  
; OTHER INFORMATION: score 3.5  
; OTHER INFORMATION: seq AUSLFYATDTSHG/SE  
US-08-905-223-168

Sequence 14033, A  
Sequence 213, App  
Sequence 76775, A  
Sequence 13453, A  
Sequence 15770, A  
Sequence 13969, A  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 16, Appl  
Sequence 72, Appl  
Sequence 37789, A  
Sequence 44271, A  
Sequence 155566,  
Sequence 155567,  
Sequence 155568,  
Sequence 155569,

```

; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 649_
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)..(295)
US-09-799-451-649

Query Match          74.4%; Score 18.6; DB 4; Length 1169;
Best Local Similarity 84.0%; Pred.No.25;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 TCAGAACCCCTGTGATGAATCAACAG 25
        ||||| ||||| ||||| |||||
Db      213 TCAGAACCATGGGATGATCAGCAG 189

RESULT 4
US-09-976-594-356/c
; Sequence 356, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 356
; LENGTH: 5520
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 198008.8
US-09-976-594-356

Query Match          74.4%; Score 18.6; DB 4; Length 5520;
Best Local Similarity 84.0%; Pred.No.35;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 TCAGAACCCCTGTGATGAATCAACAG 25
        ||||| ||||| ||||| |||||
Db      3384 TCAGAACCATGGGATGATCAGCAG 3360

RESULT 5
US-09-045-973-6
; Sequence 6, Application US/09045973
; Patent No. 6165767
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guesler, Karl J.
; APPLICANT: Baughn, Mariah
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/045,973  
FILING DATE: Filed Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0491 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1729 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRSTNOT16  
CLONE: 3041794  
US-09-045-973-6

Query Match	70.4%;	Score 17.6;	DB 3;	Length 1729;
Best Local Similarity	83.3%;	Pred. No. 83;		
Matches	20;	Conservative 0;	Mismatches 4;	Indels 0; Gaps 0;
QY	2	CAGAACCTGTGATGAATCAACG	25	
Db	166	CAGAAGCTGTGATGAATCACTAG	189	

RESULT 6  
US-09-949-016-16539  
; Sequence 16539, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 16539  
; LENGTH: 45684  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(45684)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16539

Query Match	70.4%	Score 17.6;	DB 4;	Length 45684;
Best Local Similarity	83.3%;	Pred. No. 1.7e+02;		
Matches 20;	Conservative 4;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY      2  CAGAACCTGTGATGAATCAACAG  25
      || ||||| ||||| |||||
Db     394  CATAACCTGTGATTCTTCAACAG  417

RESULT 7
US-09-751-389-3
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

```

```

Query Match      70.4%; Score 17.6; DB 4; Length 786431;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGTGATGAATCAACA 24
Db 507769 TCAGAGCTCTGTGTGATGACCAACA 507792

```

RESULT 8  
US-09-949-016-3362  
; Sequence 3362, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3362  
; LENGTH: 2234  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-3362

```

Query Match      69.6%; Score 17.4; DB 4; Length 2234;
Best Local Similarity 94.7%; Pred. NO. 1.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TCAGAACCCCTGTGATCAAT 19
         |||||
Db       35 TCAGAACCCCTGTGATCAAT 53
         |||||

RESULT 9

```

```
US-09-949-016-15104
; Sequence 15104, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15104
; LENGTH: 33908
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..-(33908)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15104

Query Match          69.6%; Score 17.4; DB 4; Length 33908;
Best Local Similarity 94.7%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TCAGAACCTGTGATCAAT 19
Db      2035 TCAGAACCTGTGATCAAT 2053

RESULT 10
US-09-949-016-109231/c
; Sequence 109231, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109231
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-109231

Query Match          68.8%; Score 17.2; DB 4; Length 601;
Best Local Similarity 86.4%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 CAGAACCTGTGATCAATCAAC 23
Db      291 CAGAACCTGTGATCAATTATC 270

RESULT 11
US-09-949-016-16430/c
; Sequence 16430, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16430
; LENGTH: 57726
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16430

Query Match          68.8%; Score 17.2; DB 4; Length 57726;
Best Local Similarity 86.4%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      4 GAACCTGTGATCAATCAACAG 25
Db      48942 GAACCTGGATGAATCAAAAG 48921

RESULT 12
US-09-949-016-14757
; Sequence 14757, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14757
; LENGTH: 154600
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14757

Query Match          68.8%; Score 17.2; DB 4; Length 154600;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 CAGAACCTGTGATCAATCAAC 23
Db      33512 CAGAACCTGTGATGAATTATC 33533

RESULT 13
US-09-799-451-489
; Sequence 489, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
```

```
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 489
; LENGTH: 6653
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(5844)
; US-09-799-451-489

Query Match      68.0%; Score 17; DB 4; Length 6653;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 TCAGAACCCCTGTGATGAATCAACAG 25
Db      1938 TCAGAACAGCTGTGATGAATCAGCAG 1962

RESULT 14
US-08-961-527-36/c
; Sequence 36, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21706 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-36

Query Match      68.0%; Score 17; DB 3; Length 21706;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 TCAGAACCCCTGTGATGAATCAACAG 25
Db      7040 TCAAAATCCCGGTATTGAATCAACAG 7016

RESULT 15
US-09-949-016-16928/c
; Sequence 16928, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16928
; LENGTH: 524032
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(524032)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-16928

Query Match      68.0%; Score 17; DB 4; Length 524032;
Best Local Similarity 80.0%; Pred. No. 5.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 TCAGAACCCCTGTGATGAATCAACAG 25
Db      346641 TCAAAATCCTGTGAGGATCAACTG 346617

Search completed: May 16, 2005, 06:11:30
Job time : 100.555 secs
```

**PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 02:00:36 ; Search time 1185.16 Seconds  
(without alignments)  
129.102 Million cell updates/sec

Title: US-10-808-187A-2472

Perfect score: 25

Sequence: 1 tcgaaccctgtgaatcaacac 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	25	100.0	25	19	US-10-808-187-2472
2	25	100.0	25	19	US-10-808-187-1
3	25	100.0	1213	19	US-10-808-187-11
4	25	100.0	5262	19	US-10-699-936-9
5	25	100.0	28920	19	US-10-889-447-5
6	25	100.0	28920	19	US-10-889-447-6
7	25	100.0	29231	19	US-10-889-447-4
8	25	100.0	29430	19	US-10-889-447-7
9	25	100.0	29727	18	US-10-839-729-15
10	25	100.0	29727	18	US-10-827-757-1
11	25	100.0	29727	19	US-10-889-447-8
					Sequence 2472, Ap
					Sequence 1, Appli
					Sequence 11, Appl
					Sequence 9, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 4, Appli
					Sequence 7, Appli
					Sequence 15, Appli
					Sequence 1, Appli
					Sequence 8, Appli

C 12	25	100.0	29727	19	US-10-699-936-1	Sequence 1, Appli
C 13	25	100.0	29736	18	US-10-839-729-17	Sequence 17, Appli
C 14	25	100.0	29736	19	US-10-889-447-9	Sequence 9, Appli
C 15	25	100.0	29736	19	US-10-699-936-3	Sequence 3, Appli
C 16	25	100.0	29742	18	US-10-839-729-16	Sequence 16, Appli
C 17	25	100.0	29742	19	US-10-808-187-15	Sequence 15, Appli
C 18	25	100.0	29742	19	US-10-808-187-16	Sequence 16, Appli
C 19	25	100.0	29742	19	US-10-808-187-240	Sequence 240, App
C 20	25	100.0	29742	19	US-10-808-187-737	Sequence 737, App
C 21	25	100.0	29742	19	US-10-808-187-1108	Sequence 1108, App
C 22	25	100.0	29742	19	US-10-808-187-1590	Sequence 1590, Ap
C 23	25	100.0	29742	19	US-10-808-187-1965	Sequence 1965, Ap
C 24	25	100.0	29742	19	US-10-889-447-10	Sequence 10, Appli
C 25	25	100.0	29751	18	US-10-839-729-14	Sequence 14, Appli
C 26	25	100.0	29751	19	US-10-856-529-1	Sequence 1, Appli
C 27	25	100.0	29751	19	US-10-626-879-67	Sequence 67, Appli
C 28	25	100.0	29751	19	US-10-889-447-1	Sequence 1, Appli
C 29	25	100.0	29751	19	US-10-889-447-2	Sequence 2, Appli
C 30	25	100.0	29751	19	US-10-699-936-2	Sequence 2, Appli
C 31	18.8	75.2	2400	16	US-10-292-408-19	Sequence 19, Appli
C 32	18.6	74.4	1169	17	US-10-302-172-649	Sequence 649, App
C 33	18.6	74.4	2114	13	US-10-027-632-97405	Sequence 97405, A
C 34	18.6	74.4	2114	13	US-10-027-632-97406	Sequence 97406, A
C 35	18.6	74.4	2114	13	US-10-027-632-97407	Sequence 97407, A
C 36	18.6	74.4	2114	17	US-10-027-632-97405	Sequence 97405, A
C 37	18.6	74.4	2114	17	US-10-027-632-97406	Sequence 97406, A
C 38	18.6	74.4	2114	17	US-10-027-632-97407	Sequence 97407, A
C 39	18.6	74.4	4511	10	US-09-764-891-7670	Sequence 7670, Ap
C 40	18.6	74.4	192992	19	US-10-461-862-95	Sequence 95, Appli
C 41	18.2	72.8	266	18	US-10-425-115-45030	Sequence 45030, A
C 42	18.2	72.8	276	10	US-09-237-183A-2245	Sequence 2245, Ap
C 43	18.2	72.8	433	17	US-10-424-599-99740	Sequence 99740, A
C 44	18.2	72.8	535	19	US-10-911-704-375	Sequence 375, App
C 45	18.2	72.8	1979	18	US-10-437-963-19059	Sequence 19059, A

```
; NUMBER OF SEQ ID NOS: 2476
; SOFTWARE: PatentIn ver. 3.2
; SEQ ID NO 2472
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-808-187-2472
```

```
Query Match          100.0%; Score 25; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
|||||
Db 1 TCAGAACCCCTGTGATGAATCAACAG 25
```

## RESULT 2

```
US-10-808-187-1/c
; Sequence 1, Application US/10808187
; Publication No. US20050090909A1
; GENERAL INFORMATION:
; APPLICANT: PEIRIS, JOSEPH S. M.
; APPLICANT: YUEN, KWOK YUNG
; APPLICANT: POON, LIT MAN
; APPLICANT: GUAN, YI
; APPLICANT: CHAN, KWOK HUNG
; APPLICANT: NICHOLLS, JOHN
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE
; FILE REFERENCE: V9661.0078
; CURRENT APPLICATION NUMBER: US/10/808,187
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 60/457,031
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: 60/457,730
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: 60/459,931
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/461,265
; PRIOR FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: 60/462,805
; PRIOR FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 60/468,139
; PRIOR FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: 60/464,886
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/471,200
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 2476
; SOFTWARE: PatentIn ver. 3.2
; SEQ ID NO 1
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Human severe acute respiratory system virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(646)
```

```
US-10-808-187-1
```

```
Query Match          100.0%; Score 25; DB 19; Length 646;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
|||||
Db 69 TCAGAACCCCTGTGATGAATCAACAG 45
```

```
RESULT 3
US-10-808-187-11/c
; Sequence 11, Application US/10808187
; Publication No. US20050090909A1
; GENERAL INFORMATION:
; APPLICANT: PEIRIS, JOSEPH S. M.
; APPLICANT: YUEN, KWOK YUNG
; APPLICANT: POON, LIT MAN
; APPLICANT: GUAN, YI
; APPLICANT: CHAN, KWOK HUNG
; APPLICANT: NICHOLLS, JOHN
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE
; FILE REFERENCE: V9661.0078
; CURRENT APPLICATION NUMBER: US/10/808,187
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 60/457,031
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: 60/457,730
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: 60/459,931
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/460,357
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/461,265
; PRIOR FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: 60/462,805
; PRIOR FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 60/468,139
; PRIOR FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: 60/464,886
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/471,200
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 2476
; SOFTWARE: PatentIn ver. 3.2
; SEQ ID NO 11
; LENGTH: 1213
; TYPE: DNA
; ORGANISM: Human severe acute respiratory system virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1213)
```

```
US-10-808-187-11
```

```
Query Match          100.0%; Score 25; DB 19; Length 1213;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
|||||
Db 636 TCAGAACCCCTGTGATGAATCAACAG 612
```

## RESULT 4

```
US-10-699-936-9/c
; Sequence 9, Application US/10699936
; Publication No. US2005009582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; PRIOR FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 5262
```



```
; TYPE: DNA
; ORGANISM: SARS coronavirus Shanghai QXC
US-10-699-936-9

Query Match      100.0%; Score 25; DB 19; Length 5262;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
Db 2051 TCAGAACCCCTGTGATGAATCAACAG 2027

RESULT 5
US-10-889-447-5/c
; Sequence 5, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ03
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-5

Query Match      100.0%; Score 25; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
Db 17512 TCAGAACCCCTGTGATGAATCAACAG 17488

RESULT 6
US-10-889-447-6/c
; Sequence 6, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ04
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-6

Query Match      100.0%; Score 25; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
Db 17512 TCAGAACCCCTGTGATGAATCAACAG 17488

RESULT 7
US-10-889-447-4/c
; Sequence 4, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 29291
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ02
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-4

Query Match      100.0%; Score 25; DB 19; Length 29291;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
Db 17838 TCAGAACCCCTGTGATGAATCAACAG 17814

RESULT 8
US-10-889-447-7/c
; Sequence 7, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 29430
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate GZ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-7

Query Match      100.0%; Score 25; DB 19; Length 29430;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25
```

Db 17497 TCAGAACCTGTGATGAATCAACAG 17473

## RESULT 9

US-10-839-729-15/c

; Sequence 15, Application US/10839729

; Publication No. US20050002953A1

; GENERAL INFORMATION:

; APPLICANT: Jens Herold

; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES

; FILE REFERENCE: BIOBANK 013A

; CURRENT APPLICATION NUMBER: US/10/839,729

; CURRENT FILING DATE: 2004-05-04

; PRIOR APPLICATION NUMBER: 60/468703

; PRIOR FILING DATE: 2003-05-06

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15

; LENGTH: 29727

; TYPE: DNA

; ORGANISM: SARS Coronavirus

US-10-839-729-15

Query Match 100.0%; Score 25; DB 18; Length 29727;

Best Local Similarity 100.0%; Pred. No. 0.2;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAACAG 25

|||||

Db 17785 TCAGAACCTGTGATGAATCAACAG 17761

## RESULT 10

US-10-827-757-1/c

; Sequence 1, Application US/10827757

; Publication No. US20050004071A1

; GENERAL INFORMATION:

; APPLICANT: Comper, Wayne

; TITLE OF INVENTION: Charged Polysaccharides Resistant To Lysosomal Degradation During

; TITLE OF INVENTION: Kidney Filtration And Renal Passage And Their Use To Treat Or

; TITLE OF INVENTION: Prevent Infection By Coronaviruses

; FILE REFERENCE: 11213-007-999

; CURRENT APPLICATION NUMBER: US/10/827,757

; CURRENT FILING DATE: 2004-04-20

; PRIOR APPLICATION NUMBER: 60/464,294

; PRIOR FILING DATE: 2003-04-21

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 29727

; TYPE: DNA

; ORGANISM: SARS-related coronavirus (Urbani strain)

US-10-827-757-1

Query Match 100.0%; Score 25; DB 18; Length 29727;

Best Local Similarity 100.0%; Pred. No. 0.2;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAACAG 25

|||||

Db 17785 TCAGAACCTGTGATGAATCAACAG 17761

## RESULT 11

US-10-889-447-8/c

; Sequence 8, Application US/10889447

; Publication No. US20050075307A1

; GENERAL INFORMATION:

; APPLICANT: Bennett, C. Frank

; APPLICANT: Jain, Ravi

; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION

; FILE REFERENCE: RTS-0685US

; CURRENT APPLICATION NUMBER: US/10/889,447

; CURRENT FILING DATE: 2004-07-12

; PRIOR APPLICATION NUMBER: 60/486,670

; PRIOR FILING DATE: 2003-07-12

; NUMBER OF SEQ ID NOS: 241

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 29727

; TYPE: DNA

; ORGANISM: SARS coronavirus Urbani

US-10-889-447-8

Query Match 100.0%; Score 25; DB 19; Length 29727;

Best Local Similarity 100.0%; Pred. No. 0.2;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAACAG 25

|||||

Db 17785 TCAGAACCTGTGATGAATCAACAG 17761

## RESULT 12

US-10-699-936-1/c

; Sequence 1, Application US/10699936

; Publication No. US20050095582A1

; GENERAL INFORMATION:

; APPLICANT: Gillim-Ross, Laura

; APPLICANT: Taylor, Jill

; APPLICANT: Scholl, David R.

; APPLICANT: Wentworth, David E.

; APPLICANT: Jollick, Joseph D.

; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory

; FILE REFERENCE: DHI-07986

; CURRENT APPLICATION NUMBER: US/10/699,936

; CURRENT FILING DATE: 2003-11-03

; NUMBER OF SEQ ID NOS: 87

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 29727

; TYPE: DNA

; ORGANISM: SARS coronavirus Urbani

US-10-699-936-1

Query Match 100.0%; Score 25; DB 19; Length 29727;

Best Local Similarity 100.0%; Pred. No. 0.2;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATGAATCAACAG 25

|||||

Db 17785 TCAGAACCTGTGATGAATCAACAG 17761

## RESULT 13

US-10-839-729-17/c

; Sequence 17, Application US/10839729

; Publication No. US20050002953A1

; GENERAL INFORMATION:

; APPLICANT: Jens Herold

; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES

; TITLE OF INVENTION: AND METHODS OF USE

; FILE REFERENCE: BIOBANK.013A

; CURRENT APPLICATION NUMBER: US/10/839,729

; CURRENT FILING DATE: 2004-05-04

; PRIOR APPLICATION NUMBER: 60/468703

; PRIOR FILING DATE: 2003-05-06

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17

; LENGTH: 29736

; TYPE: DNA

; ORGANISM: SARS Coronavirus

US-10-839-729-17

Job time : 1187.16 secs

Query Match 100.0%; Score 25; DB 18; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 0.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25  
DB 17770 TCAGAACCCCTGTGATGAATCAACAG 17746

## RESULT 14

US-10-889-447-9/c  
; Sequence 9, Application US/10889447  
; Publication No. US20050075307A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Jain, Ravi  
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
; FILE REFERENCE: RFS-0685US  
; CURRENT APPLICATION NUMBER: US/10/889,447  
; PRIOR FILING DATE: 2004-07-12  
; PRIOR APPLICATION NUMBER: 60/486,670  
; PRIOR FILING DATE: 2003-07-12  
; NUMBER OF SEQ ID NOS: 241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 29736  
; TYPE: DNA  
; ORGANISM: SARS coronavirus CUHK-W1  
US-10-889-447-9

Query Match 100.0%; Score 25; DB 19; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 0.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25  
DB 17770 TCAGAACCCCTGTGATGAATCAACAG 17746

## RESULT 15

US-10-699-936-3/c  
; Sequence 3, Application US/10699936  
; Publication No. US20050095582A1  
; GENERAL INFORMATION:  
; APPLICANT: Gillim-Ross, Laura  
; APPLICANT: Taylor, Jill  
; APPLICANT: Scholl, David R.  
; APPLICANT: Wentworth, David E.  
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory  
; FILE REFERENCE: DHI-07986  
; CURRENT APPLICATION NUMBER: US/10/699,936  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 29736  
; TYPE: DNA  
; ORGANISM: SARS coronavirus CUHK-W1  
US-10-699-936-3

Query Match 100.0%; Score 25; DB 19; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 0.2;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATGAATCAACAG 25  
DB 17770 TCAGAACCCCTGTGATGAATCAACAG 17746

Search completed: May 16, 2005, 14:33:46

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:01:20 ; Search time 2823.44 Seconds  
(without alignments)  
337.038 Million cell updates/sec

Title: US-10-808-187A-2472

Perfect score: 25  
Sequence: 1 tcagaacctgtgatcaacacag 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20.8	83.2	633	7 CF231380	CF231380 PtaC0020G
C 2	20.4	81.6	843	7 CF412109	CF412109 CH3#077_F
C 3	20.4	81.6	950	7 CF412108	CF412108 CH3#077_F
C 4	20.2	80.8	270	6 CA924034	CA924034 MTU7CL.F1
C 5	20.2	80.8	675	5 BU835678	BU835678 T07B08.P
C 6	20.2	80.8	700	8 BZ073043	BZ073043 LK18F01.B
C 7	20.2	80.8	719	7 CV230420	CV230420 WS01916.B
C 8	19.8	79.2	859	9 CNS02HTG	AL198061 Tetraodon
C 9	19.4	77.6	649	9 AG012369	AG012369 Homo sapi
C 10	19.2	76.8	245	2 AW146475	AW146475 ME000423.
C 11	19.2	76.8	360	2 AW497769	AW497769 ME000792.
C 12	19.2	76.8	470	4 BM401974	BM401974 J12G05F.S
C 13	19.2	76.8	474	8 AQ155366	AQ155366 HS 3058.A
C 14	19.2	76.8	477	8 BZ372477	BZ372477 i67h11_b
C 15	19.2	76.8	532	5 BX926690	BX926690 BX926690
C 16	19.2	76.8	736	5 CL999621	CL999621 ZMMBHFO01
C 17	19.2	76.8	832	5 BU938208	BU938208 AGENCOURT
C 18	19.2	76.8	1178	8 CC227732	CC227732 CH21-156
C 19	18.8	75.2	383	2 AW903308	AW903308 CH4-NN102
C 20	18.8	75.2	485	8 AZ601060	AZ601060 IM0419107
C 21	18.8	75.2	534	8 AZ601400	AZ601400 IM0419107
C 22	18.8	75.2	735	4 BM386263	BM386263 UI-R-CN1
C 23	18.8	75.2	759	9 AG462202	AG462202 Mus muscu
C 24	18.8	75.2	934	9 CNS043P7	AL273076 Tetraodon

25	18.8	75.2	1056	9	CL085571	CL085571	ISB1-5F20
26	18.8	75.2	1403	9	CL648225	CL648225	CH213-167
27	18.6	74.4	238	7	CV372131	CV372131	PM3-CT080
28	18.6	74.4	302	9	CE475485	CE475485	tigr-gss-
C 29	18.6	74.4	367	1	AV660306	AV660306	AV660306
C 30	18.6	74.4	367	1	AV660314	AV660314	AV660314
C 31	18.6	74.4	367	1	AV660341	AV660341	AV660341
C 32	18.6	74.4	368	1	AV660099	AV660099	AV660099
C 33	18.6	74.4	371	1	AV660041	AV660041	AV660041
C 34	18.6	74.4	372	1	AV660250	AV660250	AV660250
C 35	18.6	74.4	379	8	BH126055	BH126055	RPCI-24-2
C 36	18.6	74.4	393	9	CL898889	CL898889	abg50407.
C 37	18.6	74.4	418	9	CE254009	CE254009	tigr-gss-
C 38	18.6	74.4	428	2	BF324198	BF324198	su34901.Y
C 39	18.6	74.4	435	2	AW761184	AW761184	sl64C09.Y
C 40	18.6	74.4	450	8	AZ393609	AZ393609	IM0156D14
C 41	18.6	74.4	456	2	BE090053	BE090053	RC6-BT070
C 42	18.6	74.4	514	4	BM307189	BM307189	gak37h08.
C 43	18.6	74.4	521	6	CA163929	CA163929	SCRUR2308
C 44	18.6	74.4	535	5	BU546509	BU546509	GM880008B
C 45	18.6	74.4	559	6	CA122424	CA122424	SCULLR103

## ALIGNMENTS

RESULT 1  
LOCUS CF231380/c  
DEFINITION PtaC0020G4G0414 Poplar cDNA library from cambial zone Populus alba  
x Populus tremula cDNA 5', mRNA sequence.  
ACCESSION CF231380  
VERSION CF231380.1 GI:33450809  
KEYWORDS EST.  
SOURCE Populus alba x Populus tremula  
ORGANISM Populus alba x Populus tremula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Malpighiales; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 633)  
Dejardin,A., Leple,J.-C., Lesage-Descauses,M.-C., Costa,G. and Pilate,G.  
Expressed sequence tags from poplar wood tissues - A comparative analysis from multiple libraries  
Plant Biol. 6 (1), 55-64 (2004)  
Contact: Leple JC  
Unit of Forest improvement, Genetics and Physiology  
National Institute for Agricultural Research (INRA)  
Domaine de Limere, BP20619 ARDON, 45166 OLIVET CEDEX, FRANCE  
Tel: 33 02 38 41 78 00  
Fax: 33 02 38 41 78 79  
Email: Jean-Charles.Leple@orleans.inra.fr  
PCR Primers  
FORWARD: TriplexA 5' CTGGGAAGCGCGCCATTGTG 3'  
BACKWARD: TriplexB1 5' ATACGACTCATATTAGGCGCA 3'  
Plate: PtaC0020 row: G column: 4  
Seq primer: TriplexA 5' CTGGGAAGCGCGCCATTGTG 3'.  
Location/Qualifiers  
1..633  
/organism="Populus alba x Populus tremula"  
/mol\_type="mRNA"  
/strain="clone INRA 717-1-B4"  
/db\_xref="taxon:80863"  
/sex="female"  
/tissue\_type="cambial zone harvested on the bark side"  
/dev\_stage="3-years-old poplar trees grown in the nursery"  
/clone\_lib="Poplar cDNA library from cambial zone"  
/note="A composite cDNA library was made with mRNA isolated from opposite and tension wood tissues corresponding to the cambial zone collected on the bark side after debarking the stem. In this respect, in addition to cambium cDNA, this library also contains very young phloem and very young xylem cDNA. The sampling was

done on 3 different tilted trees grown in the nursery.  
cDNA were cloned in an oriented way into SfiI (A and B)  
restriction sites. A one-step conversion of Lambda  
Triplex2 to the corresponding pTriplex2 plasmid was done  
via site-specific recombination at loxP sites (Clontech ;  
SMART cDNA library construction kit). cDNA inserts were  
PCR amplified using flanking primers and then sequenced on  
a ABI3100 Genetic Analyser (Applied Biosystem)"

## ORIGIN

Query Match 83.2%; Score 20.8; DB 7; Length 633;  
Best Local Similarity 88.0%; Pred. No. 83;  
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TCAGAACCTGTGATGAATCAACAG 25  
- ||||| ||| ||||| |||||  
Db 572 TCAGAACCTGTGATGAATCAACAG 548

## RESULT 2

CF412109 843 bp mRNA linear EST 02-SEP-2003  
LOCUS CH3#077\_F12MR Canine heart normalized cDNA Library in pBluescript  
DEFINITION Canis familiaris cDNA clone CH3#077\_F12 3', mRNA sequence.  
ACCESSION CF412109  
VERSION CF412109.1 GI:34413355  
KEYWORDS EST.  
SOURCE Canis familiaris (dog)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
1 (bases 1 to 843)  
Yi, Y., Desai, R., Olarte, M., Henthorn, P. and George A.L.  
Expressed sequence tags from Canine heart  
Unpublished (2003)  
Other ESTs: CH3#077\_F12MF  
Contact: George AL  
Division of Genetic Medicine  
Vanderbilt University  
529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA  
Tel: 615 936 2660  
Fax: 615 936 2661  
Email: al.george@vanderbilt.edu

## REFERENCE

Yi, Y., Desai, R., Olarte, M., Henthorn, P. and George A.L.

Expressed sequence tags from Canine heart

Unpublished (2003)

## COMMENT

Location/Qualifiers  
1..843  
/organism="Canis familiaris"  
/mol\_type="mRNA"  
/db\_xref="taxon:9615"  
/clone="CH3#077\_F12"  
/tissue\_type="heart"  
/cell\_type="heart"  
/dev\_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"  
/clone\_lib="Canine heart normalized cDNA Library in pBluescript"  
/note="Organ: heart; Vector: pBluescript; Site: 1: 5' of vector NotI; Site 2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal), right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dT primed"

## FEATURES

source

Query Match 81.6%; Score 20.4; DB 7; Length 843;  
Best Local Similarity 95.5%; Pred. No. 1.3e+02;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 AGAACCTGTGATGAATCAACA 24  
||||| ||||| ||||| ||||| |||||  
Db 432 AGAACCTGTGATGAATCAACA 411

## ORIGIN

Query Match 81.6%; Score 20.4; DB 7; Length 843;  
Best Local Similarity 95.5%; Pred. No. 1.3e+02;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3 AGAACCTGTGATGAATCAACA 24  
||||| ||||| ||||| ||||| |||||

Db 813 AGAACCTGTGATGAATCAACA 834

## RESULT 3

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 4

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 5

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 6

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 7

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 8

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 9

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 10

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 11

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 12

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 14

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## Conservative

## Mismatches

## Indels

## Gaps

## QY

## Db

## RESULT 15

```

REFERENCE
AUTHORS      1 (bases 1 to 270)
              Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
              Tsai,C.-J.
TITLE        Expressed sequence tags from Aspen
JOURNAL      Unpublished (2003)
COMMENT      Contact: Tsai C-J
              Plant Biotech Research Center
              Michigan Technological University, School of Forest Resources &
              Environmental Science
              1400 Townsend Drive, Houghton, MI 49931-1295, USA
              Tel: 906 487 2914
              Fax: 906 487 2915
              Email: chtsai@mtu.edu.
FEATURES
source       Location/Qualifiers
              1..270
                /organism="Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:3693"
                /clone_lib="Aspen leaf cDNA Library"
                /note="Organ: leaf"
ORIGIN
Query Match      80.8%; Score 20.2; DB 6; Length 270;
Best Local Similarity 88.0%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATCAATCAACAG 25
    ||||| || ||||| ||||| |||||
Db 106 TCAGAAATCCAGTGTGATCAACAG 130

RESULT 5
BU835678/c
LOCUS          BU835678          675 bp mRNA linear EST 15-OCT-2002
DEFINITION    T077808 Populus apical shoot cDNA library Populus tremula x Populus
              tremuloides cDNA 5 prime, mRNA sequence.
ACCESSION     BU835678
VERSION       BU835678.1 GI:24018490
KEYWORDS      EST.
SOURCE        Populus tremula x Populus tremuloides
ORGANISM      Populus tremula x Populus tremuloides
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE     Unneberg,P., Bhalerao,R.R., Jansson,S. and Sterky,F.
AUTHORS        The poplar tree transcriptome: Analysis of expressed sequence tags
              from multiple libraries
TITLE          Unpublished (2002)
JOURNAL        Contact: BHALERAO RUPALI R.
COMMENT        Umea Plant Science Center
              Department of Plant Physiology
              University of Umea, 901 87 Umea, Sweden
              Tel: +46 90 786 5279
              Fax: +46 90 786 6676
              Email: rupali.bhalerao@plantphys.umu.se.
FEATURES
source       Location/Qualifiers
              1..675
                /organism="Populus tremula x Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:47664"
                /tissue_type="apical shoot"
                /clone_lib="Populus apical shoot cDNA library"
ORIGIN
Query Match      80.8%; Score 20.2; DB 5; Length 675;
Best Local Similarity 88.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATCAATCAACAG 25
    ||||| || ||||| ||||| |||||
Db 550 TCAGAAATCCAGTGTGATCAACAG 526

```

```

RESULT 6
BZ073043
LOCUS          BZ073043          700 bp DNA linear GSS 10-OCT-2002
DEFINITION    lkf18f01.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
              sequence.
ACCESSION     BZ073043
VERSION       BZ073043.1 GI:23691559
KEYWORDS      GSS.
SOURCE        Brassica oleracea
ORGANISM      Brassica oleracea
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE     1 (bases 1 to 700)
              Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
              Nash,W., Rabinowicz,P.D. and Wilson,R.K.
              Whole genome shotgun reads from Brassica oleracea
              Unpublished (2002)
              Contact: Richard K. Wilson
              Genome Sequencing Center
              Washington University School of Medicine
              Email: submissions@watson.wustl.edu
              Plate: lkf18 row: f column: 01
              Seq primer: -2lUPPOT forward
              Class: shotgun
              High quality sequence start: 17
              High quality sequence stop: 551.
FEATURES
source       Location/Qualifiers
              1..700
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /db_xref="taxon:3712"
                /clone_lib="B.oleracea002"
                /note="Vector: pOTw13; Whole genome shotgun library from
              flowering buds. DNA was purified from a crude nuclear
              prep using Brassica oleracea TO1000DH3 buds provided by
              Thomas Osborn at the University of Wisconsin. Genomic
              DNA was provided by Pablo Rabinowicz (CSHL) and the
              shotgun library prepared at Washington University Genome
              Sequencing Center."

```

## ORIGIN

```

Query Match      80.8%; Score 20.2; DB 8; Length 700;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCAGAACCCCTGTGATCAATCAACAG 25
    ||||| || ||||| ||||| |||||
Db 253 TCAGAACCCGCGTGTGATGATAAACAG 277

```

## RESULT 7

```

CV230420
LOCUS          CV230420          719 bp mRNA linear EST 21-SEP-2004
DEFINITION    WS01916.B21.H19 PT-DX-N-A-10 Populus balsamifera subsp. trichocarpa
              cDNA clone WS01916_H19 3', mRNA sequence.
ACCESSION     CV230420
VERSION       CV230420.1 GI:52383897
KEYWORDS      EST.
SOURCE        Populus balsamifera subsp. trichocarpa (Populus trichocarpa)
ORGANISM      Populus balsamifera subsp. trichocarpa
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE     1 (bases 1 to 719)
              Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
              Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
              Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
              Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
              Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
              Bohlmann,J.
              The poplar transcriptome: Analysis of expressed sequence tags from

```

## TITLE

JOURNAL  
COMMENT  
multiple cDNA libraries  
Unpublished (2004)  
Contact: Joerg Bohlmann  
Genome BC forest genomics program  
University of British Columbia  
UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,  
Vancouver, British Columbia, Canada, V6T 1Z3  
Tel: 1-604-822-0282  
Fax: 1-604-822-6097  
Email: bohlmann@interchange.ubc.ca  
Plate: WS01916 row: H column: 19  
High quality sequence stop: 719  
POLYA=Yes.

## FEATURES

source

1. 719  
Location/Qualifiers  
/organism="Populus balsamifera subsp. trichocarpa"  
/mol\_type="mRNA"  
/cultivar="VT-125"  
/sub\_species="trichocarpa"  
/db\_xref="taxon:3694"  
/clone="WS01916 H19"  
/sex="Not determined"  
/lab\_host="E. coli DH10B T1 phage resistant cells"  
/clone\_lib="PT-DX-N-A-10"  
/note="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Outer xylem from 5 year old trees harvested every two weeks between April and October of 2002 at the University of British Columbia south campus farm in Vancouver, British Columbia. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

## ORIGIN

Query Match 80.8%; Score 20.2; DB 7; Length 719;  
Best Local Similarity 88.0%; Pred. No. 1.6e+02;  
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCAGAACCTGTGATCAATCAACAG 25  
|||||  
DB 405 TCAGATCCAGTGCTGAATCAACAG 429

## RESULT 8

CNS02HTG  
LOCUS  
DEFINITION  
Tetraodon nigroviridis genome survey sequence PUC-ori end of clone 140A04 of library G from Tetraodon nigroviridis, genomic survey sequence.  
ACCESSION  
AL198061.1 GI:7836212  
VERSION  
GSS; genome survey sequence.  
KEYWORDS  
Tetraodon nigroviridis  
SOURCE  
ORGANISM  
Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthopterygii; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

## REFERENCE

1  
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.  
Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

## COMMENT

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetraodon.

## FEATURES

source

1. 859  
Location/Qualifiers  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="140A04"  
/clone\_lib="G"  
/note="Genoscope sequence ID : COAG140BA02SP1-end : PUC-ori"

## ORIGIN

Query Match 79.2%; Score 19.8; DB 9; Length 859;  
Best Local Similarity 91.3%; Pred. No. 2.6e+02;  
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 CAGAACCTGTGATCAATCAACA 24  
|||||  
DB 575 CAGACCCCTGTGATCAATCAACA 597

## RESULT 9

AG012369/c  
LOCUS  
DEFINITION  
Homo sapiens genomic DNA, 21q region, clone: B2697SPN28, genomic survey sequence.

ACCESSION  
AG012369.1 GI:3413638  
VERSION  
GSS.  
KEYWORDS  
Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 649)

## REFERENCE

1  
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.  
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.  
Published Only in DataBase (1998)  
2 (bases 1 to 649)  
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.  
Direct Submission  
Submitted (12-AUG-1998) Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)

## FEATURES

source

1. 649  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q"  
/clone="B2697SPN28"



## ORIGIN

Query Match 77.6%; Score 19.4; DB 9; Length 649;  
 Best Local Similarity 95.2%; Pred. No. 3.8e+02;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAA 22  
 |||||  
 Db 549 CAGAACCTGTGACGAATCAA 529

RESULT 10  
 AW146475  
 LOCUS  
 DEFINITION M5000423.FER Egg stage cDNA expression library in Lambda ZAPII  
 Schistosoma mansoni cDNA 5' similar to EST A1740431 Schistosoma  
 mansoni, egg, mRNA sequence.

ACCESSION AW146475

VERSION AW146475.1 GI:6194382

KEYWORDS EST.

SOURCE Schistosoma mansoni

ORGANISM Schistosoma mansoni

REFERENCE Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
 Strigeidida; Schistosomatoidea; Schistosomidae; Schistosoma.

AUTHORS Oliveira,G.C. and Baba,J.

TITLE Cataloging Schistosoma mansoni genes with expressed sequence tags

JOURNAL Unpublished (1998)

COMMENT Contact: Oliveira, Guilherme

Lab. Parasitologia Cel. e Mol.

Centro de Pesquisas Rene Rachou - FIOCRUZ

Av. Augusto de Lima 1715 Barro Preto, Belo Horizonte, MG, CEP

30190, Brazil

Tel: 55 31 2953566

Fax: 55 31 2952115

Email: oliveira@netra.cpqrr.fiocruz.br

Insert Length: 245 Std Error: 0.00

Seq primer: M13 Reverse Universal Sequencing primer.

Location/Qualifiers

1. .245

/organism="Schistosoma mansoni"

/mol\_type="mRNA"

/strain="LE"

/db\_xref="taxon:6183"

/sex="Mixed"

/dev\_stage="Egg"

/clone\_lib="Egg stage cDNA expression library in Lambda

ZAPII"

/note="Vector: Uni-Zap XR vector, Stratagene (pBluescript

SK); Site 1: EcoRI; Site 2: XhoI; mRNA was extracted from

eggs and the library was constructed and excised

according to the manufacturer's instructions."

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 245;

Best Local Similarity 87.5%; Pred. No. 4e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 29 CAGAACCTTTTGATCAACAG 52

|||||

RESULT 11

AW497769

LOCUS

DEFINITION M5000792.FER Egg stage cDNA expression library in Lambda ZAPII

Schistosoma mansoni cDNA 5' similar to EST AW146475 Egg stage cDNA

expression library, mRNA sequence.

ACCESSION AW497769

VERSION AW497769.1 GI:7119504

KEYWORDS EST.

SOURCE Schistosoma mansoni

REFERENCE Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;

Strigeidida; Schistosomatoidea; Schistosomidae; Schistosoma.

AUTHORS Oliveira,G.C. and Baba,J.

TITLE Cataloging Schistosoma mansoni genes with expressed sequence tags

JOURNAL Unpublished (1998)

COMMENT Contact: Oliveira, Guilherme

Lab. Parasitologia Cel. e Mol.

Centro de Pesquisas Rene Rachou - FIOCRUZ

Av. Augusto de Lima 1715 Barro Preto, Belo Horizonte, MG, CEP

30190, Brazil

Tel: 55 31 2953566

Fax: 55 31 2952115

Email: oliveira@netra.cpqrr.fiocruz.br

Insert Length: 360 Std Error: 0.00

Seq primer: M13 Reverse Universal Sequencing primer.

Location/Qualifiers

1. .360

/organism="Schistosoma mansoni"

/mol\_type="mRNA"

/strain="LE"

/db\_xref="taxon:6183"

/sex="Mixed"

/dev\_stage="Egg"

/clone\_lib="Egg stage cDNA expression library in Lambda

ZAPII"

/note="Vector: Uni-Zap XR vector, Stratagene (pBluescript

SK); Site 1: EcoRI; Site 2: XhoI; mRNA was extracted from

eggs and the library was constructed and excised

according to the manufacturer's instructions."

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

RESULT 12

BM401974/c

LOCUS

DEFINITION JL2G05F Snake Bothrops insularis library IL3 Bothrops insularis

cDNA 5' mRNA sequence.

ACCESSION BM401974

VERSION BM401974.1 GI:20376603

KEYWORDS EST.

SOURCE Bothrops insularis (Island Jararaca)

ORGANISM Bothrops insularis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;

Viperidae; Crotalinae; Bothrops.

AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.

TITLE A survey of gene expression and diversity in the venom glands of

the pitviper snake Bothrops insularis through the generation of

expressed sequence tags (ESTs)

JOURNAL Gene 299 (1-2), 279-291 (2002)

MEDLINE 22347338

PUBMED 12459276

COMMENT Contact: Paulo Lee Ho

Centro de Biotecnologia

Instituto Butantan

Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900

Tel: 55 11 37 26 7222 ext. 2083

Fax: 55 11 37 26 1505

Email: hoplee@usp.br

This EST corresponds to cluster BINI75A (see Reference)

Seq primer: M13F.

Location/Qualifiers

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

RESULT 12

BM401974/c

LOCUS

DEFINITION JL2G05F Snake Bothrops insularis library IL3 Bothrops insularis

cDNA 5' mRNA sequence.

ACCESSION BM401974

VERSION BM401974.1 GI:20376603

KEYWORDS EST.

SOURCE Bothrops insularis (Island Jararaca)

ORGANISM Bothrops insularis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;

Viperidae; Crotalinae; Bothrops.

AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.

TITLE A survey of gene expression and diversity in the venom glands of

the pitviper snake Bothrops insularis through the generation of

expressed sequence tags (ESTs)

JOURNAL Gene 299 (1-2), 279-291 (2002)

MEDLINE 22347338

PUBMED 12459276

COMMENT Contact: Paulo Lee Ho

Centro de Biotecnologia

Instituto Butantan

Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900

Tel: 55 11 37 26 7222 ext. 2083

Fax: 55 11 37 26 1505

Email: hoplee@usp.br

This EST corresponds to cluster BINI75A (see Reference)

Seq primer: M13F.

Location/Qualifiers

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

Best Local Similarity 87.5%; Pred. No. 4.3e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAGAACCTGTGATCAACAG 25

|||||

Db 194 CAGAACCTTTTGATCAACAG 217

|||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 360;

## ORIGIN

ACCESSION BX926690  
VERSION BX926690.1 GI:41143538  
KEYWORDS EST.  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
1 (bases 1 to 532)  
Bonnet,A., Tosser-Klopp,G., Benne,F., Cabau,C., Villegier,S.,  
Soares,M., Bonaldo,F. and Hatey,F.  
TITLE A Pig Normalised Multi-Tissue cDNA Library  
JOURNAL Unpublished (2003)  
COMMENT Contact: Tosser-Klopp G  
Genetique Animale  
Institut National de la Recherche Agronomique  
Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan  
cedex, FRANCE  
Tel: 33 (0) 5.61.28.51.14  
Fax: 33 (0) 5.61.28.53.08  
Email: tosser@toulouse.inra.fr  
Clone distribution: AGENAE Resource centre. Francois PIUMI,  
Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du  
genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex,  
FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73  
Sequence cleaned of vector, adaptor and repetitions. Contact us  
at sigenasupport@jouy.inra.fr to obtain the chromatogram of this  
sequence.  
Plate: 0016 row: h column: 16.  
FEATURES  
source Location/Qualifiers  
1..532  
/organism="Sus scrofa"  
/mol\_type="mrna"  
/db\_xref="taxon:9823"  
/clone="scan0016d.h.16"  
/tissue\_type="mixed"  
/dev\_stage="from embryos to adults"  
/clone\_lib="Sus Scrofa library (scan)"  
/note="tissues: adipose tissue, brain, kidney, liver,  
muscle, ovary, testis, heart, hypothalamus, pancreas,  
skin, spleen, thymus, placenta, pituitary gland, seminal  
vesicle, small intestine, uterus, adrenals, bulbo urethral  
gland, cerebral trunk, epididymis, female gonad,  
gall-bladder, hippocampus, large intestine, male gonad,  
melanocytes, stomach, udder"

Query Match 76.8%; Score 19.2; DB 5; Length 532;  
Best Local Similarity 87.5%; Pred. No. 4.6e+02;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 CAGAACCTGTGATGATCAACAG 25  
||||| ||||| ||||| ||||| |||||  
Db 215 CAGAACCTGAGATGACTCAACAG 238

Search completed: May 16, 2005, 06:03:20  
Job time : 2831.44 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 21:16:45 ; Search time 392.5 Seconds  
(without alignments)  
2026.885 Million cell updates/sec

Title: US-10-808-187A-2473

Perfect score: 16

Sequence: 1 tctgcgtaggcaatcc 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16	100.0	626	14	AY451928 SARS coro
C 2	16	100.0	626	14	AY451929 SARS coro
C 3	16	100.0	626	14	AY451930 SARS coro
C 4	16	100.0	626	14	AY451931 SARS coro
C 5	16	100.0	626	14	AY451932 SARS coro
C 6	16	100.0	626	14	AY451933 SARS coro
C 7	16	100.0	626	14	AY451934 SARS coro
C 8	16	100.0	626	14	AY451935 SARS coro
C 9	16	100.0	626	14	AY451936 SARS coro
C 10	16	100.0	626	14	AY451937 SARS coro
C 11	16	100.0	626	14	AY451938 SARS coro
C 12	16	100.0	626	14	AY451939 SARS coro
C 13	16	100.0	626	14	AY451940 SARS coro
C 14	16	100.0	626	14	AY451941 SARS coro
C 15	16	100.0	626	14	AY451942 SARS coro
C 16	16	100.0	626	14	AY451943 SARS coro
C 17	16	100.0	626	14	AY451944 SARS coro
C 18	16	100.0	626	14	AY451945 SARS coro
C 19	16	100.0	646	14	AY268070 SARS coro

C 20	16	100.0	677	14	AY443086S05
C 21	16	100.0	6067	14	AY534762S3
C 22	16	100.0	8796	14	AY534758S2
C 23	16	100.0	13471	14	AY304490
C 24	16	100.0	13471	14	AY304492
C 25	16	100.0	26333	14	AY286320
C 26	16	100.0	29013	14	AY463060
C 27	16	100.0	29350	14	AY394999
C 28	16	100.0	29350	14	AY395000
C 29	16	100.0	29350	14	AY395001
C 30	16	100.0	29350	14	AY395002
C 31	16	100.0	29433	14	AY394977
C 32	16	100.0	29530	14	AY394985
C 33	16	100.0	29573	14	AY338174
C 34	16	100.0	29573	14	AY338175
C 35	16	100.0	29573	14	AY348314
C 36	16	100.0	29577	14	AY559094
C 37	16	100.0	29592	14	AY463059
C 38	16	100.0	29620	14	AY395004
C 39	16	100.0	29640	14	AY394978
C 40	16	100.0	29645	14	AY394979
C 41	16	100.0	29646	14	AY394982
C 42	16	100.0	29647	14	AY395003
C 43	16	100.0	29661	14	AY559086
C 44	16	100.0	29665	14	AY394988
C 45	16	100.0	29670	14	AY559082

#### ALIGNMENTS

RESULT 1  
AY451928/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

ORIGIN

AY451928 626 bp RNA linear VRL 28-FEB-2004  
SARS coronavirus TW-HPI isolate TW-HPI\_SCI8 replicase 1B gene,  
partial cds.  
AY451928  
AY451928.1 GI:42741328  
SARS coronavirus TW-HPI  
SARS coronavirus TW-HPI  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.  
1 (bases 1 to 626)  
Lan, Y.-C., Chen, H.-Y., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,  
Yang, J.-Y., Chen, H.-Y., and Chen, Y.-M. Arthur.  
Molecular Epidemiology of SARS in Taiwan - Development of a Simple  
Method for Tracing the Origin and Dissemination of SARS  
Unpublished  
2 (bases 1 to 626)  
Lan, Y.-C., Chen, H.-Y., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J.,  
Yang, J.-Y., Chen, H.-Y., and Chen, Y.-M. Arthur.  
Direct Submission  
Submitted (28-OCT-2003) National Yang-Ming University, AIDS  
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,  
Taipei, Taiwan 112, R.O.C.  
Location/Qualifiers  
1. .626  
/organism="SARS coronavirus TW-HPI"  
/mol\_type="genomic RNA"  
/isolate="TW-HPI\_SCI8"  
/db\_xref="taxon:264371"  
<1..>626  
/codon\_start=1  
/product="replicase 1B"  
/protein\_id="AAS44814.1"  
/db\_xref="GI:42741329"  
/translation="ISMATNYDLSSVNNARLAKHYVIGDPAQLPAPRTLLTKTLEP  
EYVNSVRLMKITIGPDWFLGTCRCPAEIVDTVSALVDYDNKLKAHKQKSAQCFPMFYK  
GWITHDVSSAINEPQIGVWREFLTRNPANKEAVFISPNSONAVASLIGLPTQTVDS  
SQSGSEYDYVIFTQTTFATSCNVRNFRVAITRAKIGILCIMSRLDLYD"

Query Match	100.0%;	Score 16;	DB 14;	Length 626;	
Best Local Similarity	100.0%;	Pred. No. 1.3e+02;			
Matches	16;	Conservative	0;	Mismatches	0;
Indels			0;	Gaps	
QY	1	TCTGCGTAGGCAATCC	16		
DB	469	TCTGCGTAGGCAATCC	454		
RESULT 2					
AY451929/c					
LOCUS	AY451929	SARS coronavirus TW-HP2 isolate TW-HP2_SC18	626 bp	RNA	linear
DEFINITION		partial cds.			VRL 28-FEB-2004
ACCESSION	AY451929	1	GI:42741330		
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					

```
/db_xref="GI:42741335"
/translations="ISMATNYDLVSVNARLRAKHVYVIGDPAQLPAPRILLTKGTLEP
EYFNSVCRMLKTIQPMFLGTGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCFMFYK
GVITHDVSSAINRPQIGVREFLTRNPARKAVFISPYNSQNAVASKILGLPTQTVD
SOGSEYDYVIFTQTETAHSCNVNRFNVAITRAKIGILCIMSDDRLDYD"
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCTGCGTAGGCAATCC 16
|||||
Db 469 TCTGCGTAGGCAATCC 454

RESULT 5
AY451932/c
LOCUS SARS coronavirus TW-JC2 isolate TW-JC2_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451932
VERSION AY451932.1 GI:42741336
KEYWORDS
SOURCE SARS coronavirus TW-JC2
ORGANISM SARS coronavirus TW-JC2
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1..626
/organism="SARS coronavirus TW-JC2"
/mol_type="genomic RNA"
/isolate="TW-JC2 SC18"
/db_xref="taxon:264375"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44818.1"
/db_xref="GI:42741337"
/translations="ISMATNYDLVSVNARLRAKHVYVIGDPAQLPAPRILLTKGTLEP
EYFNSVCRMLKTIQPMFLGTGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCFMFYK
GVITHDVSSAINRPQIGVREFLTRNPARKAVFISPYNSQNAVASKILGLPTQTVD
SOGSEYDYVIFTQTETAHSCNVNRFNVAITRAKIGILCIMSDDRLDYD"
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCTGCGTAGGCAATCC 16
|||||
Db 469 TCTGCGTAGGCAATCC 454

RESULT 6
AY451933/c
LOCUS SARS coronavirus TW-KC1 isolate TW-KC1_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451933
```

```
AY451933.1 GI:42741338
SARS coronavirus TW-KC1
SARS coronavirus TW-KC1
Viruses; sRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1..626
/organism="SARS coronavirus TW-KC1"
/mol_type="genomic RNA"
/isolate="TW-KC1 SC18"
/db_xref="taxon:264376"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44819.1"
/db_xref="GI:42741339"
/translations="ISMATNYDLVSVNARLRAKHVYVIGDPAQLPAPRILLTKGTLEP
EYFNSVCRMLKTIQPMFLGTGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCFMFYK
GVITHDVSSAINRPQIGVREFLTRNPARKAVFISPYNSQNAVASKILGLPTQTVD
SOGSEYDYVIFTQTETAHSCNVNRFNVAITRAKIGILCIMSDDRLDYD"
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCTGCGTAGGCAATCC 16
|||||
Db 469 TCTGCGTAGGCAATCC 454

RESULT 7
AY451934/c
LOCUS SARS coronavirus TW-KC3 isolate TW-KC3_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451934
VERSION AY451934.1 GI:42741340
KEYWORDS
SOURCE SARS coronavirus TW-KC3
ORGANISM SARS coronavirus TW-KC3
REFERENCE 1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.
FEATURES
source
1..626
/organism="SARS coronavirus TW-KC3"
```

```

/mol_type="genomic RNA"
/isolate="TW-KC3_SC18"
/db_xref="taxon:264377"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44820.1"
/db_xref="GI:42741341"
/translation="ISMATNYDLSVNNARLRAKHVYVIGDPAQLPAPRTLTKGTLEP
EYFNSVCRMLKTIQDMFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCFKMFYK
GVITHDVSSAINRPOIGVVRFLNPAWRKAVFISPYNSQNAVASKILGLPTQTVDS
SQGSEYDYVIFTQTETAHSCNVNRFNVAITRAKIGILCINSDRLDYD"

```

## CDS

## ORIGIN

```

Query Match      100.0%; Score 16; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
|||||
DB 469 TCTGCGTAGGCAATCC 454

```

```

RESULT 8
AY451935/c
LOCUS SARS coronavirus TW-GD1 isolate TW-GD1_SC18 replicase 1B gene,
partial cds.
ACCESSION AY451935
VERSION AY451935.1 GI:42741342
KEYWORDS
SOURCE
ORGANISM SARS coronavirus TW-GD1
SARS coronavirus TW-GD1
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.

```

## FEATURES

## source

```

1..626
/organism="SARS coronavirus TW-GD1"
/mol_type="genomic RNA"
/isolate="TW-GD1_SC18"
/db_xref="taxon:264378"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44821.1"
/db_xref="GI:42741343"
/translation="ISMATNYDLSVNNARLRAKHVYVIGDPAQLPAPRTLTKGTLEP
EYFNSVCRMLKTIQDMFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCFKMFYK
GVITHDVSSAINRPOIGVVRFLNPAWRKAVFISPYNSQNAVASKILGLPTQTVDS
SQGSEYDYVIFTQTETAHSCNVNRFNVAITRAKIGILCINSDRLDYD"

```

## CDS

## ORIGIN

```

Query Match      100.0%; Score 16; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
|||||
DB 469 TCTGCGTAGGCAATCC 454

```

## RESULT 9

## LOCUS

```

AY451936/c
LOCUS SARS coronavirus TW-GD2 isolate TW-GD2_SC18 replicase 1B gene,
partial cds.
ACCESSION AY451936
VERSION AY451936.1 GI:42741344
KEYWORDS
SOURCE
ORGANISM SARS coronavirus TW-GD2
SARS coronavirus TW-GD2
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2003) National Yang-Ming University, AIDS
Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
Taipei, Taiwan 112, R.O.C.

```

## FEATURES

## source

```

1..626
/organism="SARS coronavirus TW-GD2"
/mol_type="genomic RNA"
/isolate="TW-GD2_SC18"
/db_xref="taxon:264379"
<1..>626
/codon_start=1
/product="replicase 1B"
/protein_id="AAS44822.1"
/db_xref="GI:42741345"
/translation="ISMATNYDLSVNNARLRAKHVYVIGDPAQLPAPRTLTKGTLEP
EYFNSVCRMLKTIQDMFLGTCRCPAEIVDTVSALVYDNKLAHKDKSAQCFKMFYK
GVITHDVSSAINRPOIGVVRFLNPAWRKAVFISPYNSQNAVASKILGLPTQTVDS
SQGSEYDYVIFTQTETAHSCNVNRFNVAITRAKIGILCINSDRLDYD"

```

## CDS

## ORIGIN

```

Query Match      100.0%; Score 16; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
|||||
DB 469 TCTGCGTAGGCAATCC 454

```

## RESULT 10

## LOCUS

```

AY451937/c
LOCUS SARS coronavirus TW-GD3 isolate TW-GD3_SC18 replicase 1B gene,
partial cds.
ACCESSION AY451937
VERSION AY451937.1 GI:42741346
KEYWORDS
SOURCE
ORGANISM SARS coronavirus TW-GD3
SARS coronavirus TW-GD3
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
1 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.
TITLE Molecular Epidemiology of SARS in Taiwan - Development of a Simple
Method for Tracing the Origin and Dissemination of SARS
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 626)
AUTHORS Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
Yang,J.-Y., Chen,H.-Y. and Chen,Y.-M.Arthur.

```



```

TITLE      Direct Submission
JOURNAL    Submitted (28-OCT-2003) National Yang-Ming University, AIDS
           Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
           Taipei, Taiwan 112, R.O.C.
FEATURES   source
           1. .626
           /organism="SARS coronavirus TW-GD3"
           /mol_type="genomic RNA"
           /isolate="TW-GD3_SC18"
           /db_xref="taxon:264380"
           <1. .>.626
           /codon_start=1
           /product="replicase 1B"
           /protein_id="AA544823.1"
           /db_xref="GI:42741347"
           /translations="ISMATNYDLSVVNARLRAKHVYIGDPAQLPAPRTLLTKGTLEP
           EYFNSVRLMKITIGPMFLGTCRCRPAEIVDTVSALVYDNKLKAHKDKSAQCQCFMFYK
           GVITHDVSSAINRPOIGVREFLTRNPAMKAVFISPYNSONAVASKILGLPTQTVD
           SQSEYDYVIFTQTTHSCNVNRNVAITRAKIGILCIMSDDRLYD"

ORIGIN
Query Match      100.0%; Score 16; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCTGCGTAGGCAATCC 16
       |||||
Db      469 TCTGCGTAGGCAATCC 454

RESULT 11
AY451938/c
LOCUS   SARS coronavirus TW-GD4 isolate TW-GD4_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451938
VERSION   AY451938.1 GI:42741348
KEYWORDS
SOURCE    SARS coronavirus TW-GD4
ORGANISM  SARS coronavirus TW-GD4
          Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
          Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS   Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
          Yang,J.-Y., Chen,H.-Y., and Chen,Y.-M.Arthur.
TITLE     Molecular Epidemiology of SARS in Taiwan - Development of a Simple
          Method for Tracing the Origin and Dissemination of SARS
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS   Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
          Yang,J.-Y., Chen,H.-Y., and Chen,Y.-M.Arthur.
TITLE     Direct Submission
JOURNAL   Submitted (28-OCT-2003) National Yang-Ming University, AIDS
          Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
          Taipei, Taiwan 112, R.O.C.
FEATURES   source
           1. .626
           /organism="SARS coronavirus TW-GD4"
           /mol_type="genomic RNA"
           /isolate="TW-GD4_SC18"
           /db_xref="taxon:264381"
           <1. .>.626
           /codon_start=1
           /product="replicase 1B"
           /protein_id="AA544824.1"
           /db_xref="GI:42741349"
           /translations="ISMATNYDLSVVNARLRAKHVYIGDPAQLPAPRTLLTKGTLEP
           EYFNSVRLMKITIGPMFLGTCRCRPAEIVDTVSALVYDNKLKAHKDKSAQCQCFMFYK
           GVITHDVSSAINRPOIGVREFLTRNPAMKAVFISPYNSONAVASKILGLPTQTVD
           SQSEYDYVIFTQTTHSCNVNRNVAITRAKIGILCIMSDDRLYD"

ORIGIN
Query Match      100.0%; Score 16; DB 14; Length 626;

```

```

Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCTGCGTAGGCAATCC 16
       |||||
Db      469 TCTGCGTAGGCAATCC 454

RESULT 12
AY451939/c
LOCUS   SARS coronavirus TW-GD5 isolate TW-GD5_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451939
VERSION   AY451939.1 GI:42741350
KEYWORDS
SOURCE    SARS coronavirus TW-GD5
ORGANISM  SARS coronavirus TW-GD5
          Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
          Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS   Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
          Yang,J.-Y., Chen,H.-Y., and Chen,Y.-M.Arthur.
TITLE     Molecular Epidemiology of SARS in Taiwan - Development of a Simple
          Method for Tracing the Origin and Dissemination of SARS
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 626)
AUTHORS   Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,
          Yang,J.-Y., Chen,H.-Y., and Chen,Y.-M.Arthur.
TITLE     Direct Submission
JOURNAL   Submitted (28-OCT-2003) National Yang-Ming University, AIDS
          Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou,
          Taipei, Taiwan 112, R.O.C.
FEATURES   source
           1. .626
           /organism="SARS coronavirus TW-GD5"
           /mol_type="genomic RNA"
           /isolate="TW-GD5_SC18"
           /db_xref="taxon:264382"
           <1. .>.626
           /codon_start=1
           /product="replicase 1B"
           /protein_id="AA544825.1"
           /db_xref="GI:42741351"
           /translations="ISMATNYDLSVVNARLRAKHVYIGDPAQLPAPRTLLTKGTLEP
           EYFNSVRLMKITIGPMFLGTCRCRPAEIVDTVSALVYDNKLKAHKDKSAQCQCFMFYK
           GVITHDVSSAINRPOIGVREFLTRNPAMKAVFISPYNSONAVASKILGLPTQTVD
           SQSEYDYVIFTQTTHSCNVNRNVAITRAKIGILCIMSDDRLYD"

ORIGIN
Query Match      100.0%; Score 16; DB 14; Length 626;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCTGCGTAGGCAATCC 16
       |||||
Db      469 TCTGCGTAGGCAATCC 454

RESULT 13
AY451940/c
LOCUS   SARS coronavirus TW-YM1 isolate TW-YM1_SC18 replicase 1B gene,
DEFINITION partial cds.
ACCESSION AY451940
VERSION   AY451940.1 GI:42741352
KEYWORDS
SOURCE    SARS coronavirus TW-YM1
ORGANISM  SARS coronavirus TW-YM1
          Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
          Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 626)
AUTHORS   Lan,Y.-C., Chen,Y.-J., Lee,C.-M., Liu,T.-T., Lu,J.-J., Chan,Y.-J.,

```

**TITLE** Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.  
Molecular Epidemiology of SARS in Taiwan - Development of a Simple Method for Tracing the Origin and Dissemination of SARS

**JOURNAL****REFERENCE**

2 (bases 1 to 626)  
Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.

**TITLE**

Direct Submission  
Submitted (28-OCT-2003) National Yang-Ming University, AIDS Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou, Taipei, Taiwan 112, R.O.C.

**FEATURES****source**

1..626  
Location/Qualifiers

/organism="SARS coronavirus TW-YM1"

/mol\_type="genomic RNA"

/isolate="TW-YM1\_SC18"

/db\_xref="taxon:264383"

<1..>626

/codon\_start=1

/product="replicase 1B"

/protein\_id="AA544826.1"

/db\_xref="GI:42741353"

/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLTKGTLEP  
EYFNSVCRMLKTIQDMPFGTCRCRCPAEIVDTVSALVDNKLKAHKDKSAOCFKMFYK  
GVITHDVSSAINRPOIGVVRFLTRNPARKAVFISPSYNSQNAVASKILGLPTQTVD  
SQGSEYDYVIFTQTETAHSCNVNRFNVAITRAKIGILCINMSDRDLVD"

**ORIGIN**

Query Match 100.0%; Score 16; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCGTAGGCAATCC 16

|||||

Db 469 TCTGCGTAGGCAATCC 454

**RESULT 14**

AY451941/c

LOCUS

DEFINITION SARS coronavirus TW-YM2 isolate TW-YM2\_SC18 replicase 1B gene,  
partial cds.

ACCESSION

AY451941

VERSION

AY451941.1 GI:42741354

KEYWORDS

SOURCE

ORGANISM

SARS coronavirus TW-YM2

Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

Coronaviridae; Coronavirus.

REFERENCE

AUTHORS

Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.

TITLE

Molecular Epidemiology of SARS in Taiwan - Development of a Simple Method for Tracing the Origin and Dissemination of SARS

Unpublished

2 (bases 1 to 626)

Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.

Direct Submission

Submitted (28-OCT-2003) National Yang-Ming University, AIDS Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou, Taipei, Taiwan 112, R.O.C.

FEATURES

source

1..626

/organism="SARS coronavirus TW-YM2"

/mol\_type="genomic RNA"

/isolate="TW-YM2\_SC18"

/db\_xref="taxon:264384"

<1..>626

/codon\_start=1

/product="replicase 1B"

/protein\_id="AA544827.1"

/db\_xref="GI:42741355"

**CDS**

/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLTKGTLEP  
EYFNSVCRMLKTIQDMPFGTCRCRCPAEIVDTVSALVDNKLKAHKDKSAOCFKMFYK  
GVITHDVSSAINRPOIGVVRFLTRNPARKAVFISPSYNSQNAVASKILGLPTQTVD  
SQGSEYDYVIFTQTETAHSCNVNRFNVAITRAKIGILCINMSDRDLVD"

**ORIGIN**

Query Match 100.0%; Score 16; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCGTAGGCAATCC 16

|||||

Db 469 TCTGCGTAGGCAATCC 454

**RESULT 15**

AY451942/c

LOCUS

DEFINITION SARS coronavirus TW-YM3 isolate TW-YM3\_SC18 replicase 1B gene,  
partial cds.

ACCESSION

AY451942

VERSION

AY451942.1 GI:42741356

KEYWORDS

SOURCE

ORGANISM

SARS coronavirus TW-YM3

Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

Coronaviridae; Coronavirus.

REFERENCE

AUTHORS

Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.

TITLE

Molecular Epidemiology of SARS in Taiwan - Development of a Simple Method for Tracing the Origin and Dissemination of SARS

Unpublished

2 (bases 1 to 626)

Lan, Y.-C., Chen, Y.-J., Lee, C.-M., Liu, T.-T., Lu, J.-J., Chan, Y.-J., Yang, J.-Y., Chen, H.-Y. and Chen, Y.-M. Arthur.

Direct Submission

Submitted (28-OCT-2003) National Yang-Ming University, AIDS Prevention and Research Center, 155, Li-Nong St., Sec.2, Peitou, Taipei, Taiwan 112, R.O.C.

FEATURES

source

1..626

/organism="SARS coronavirus TW-YM3"

/mol\_type="genomic RNA"

/isolate="TW-YM3\_SC18"

/db\_xref="taxon:264385"

<1..>626

/codon\_start=1

/product="replicase 1B"

/protein\_id="AA544828.1"

/db\_xref="GI:42741357"

/translation="ISMATNYDLSVNNARLRAKHVYIGDPAQLPAPRTLTKGTLEP  
EYFNSVCRMLKTIQDMPFGTCRCRCPAEIVDTVSALVDNKLKAHKDKSAOCFKMFYK  
GVITHDVSSAINRPOIGVVRFLTRNPARKAVFISPSYNSQNAVASKILGLPTQTVD  
SQGSEYDYVIFTQTETAHSCNVNRFNVAITRAKIGILCINMSDRDLVD"

**ORIGIN**

Query Match 100.0%; Score 16; DB 14; Length 626;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCGTAGGCAATCC 16

|||||

Db 469 TCTGCGTAGGCAATCC 454

Search completed: May 16, 2005, 02:00:24

Job time : 383.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 17:17:00 ; Search time 205.25 Seconds  
(without alignments)  
461.466 Million cell updates/sec

Title: US-10-808-187A-2473

Perfect score: 16

Sequence: 1 tctgcgtaggcaatcc 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseqn16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16	100.0	29751	12	ADJ39000 SARS coro
C 2	15	93.8	875	8	Abz51221 Aspergill
C 3	14.4	90.0	746	6	Abv31202 E. coli J
C 4	14.4	90.0	801	6	Abz13359 Arabidops
C 5	14.4	90.0	1024	10	Abz83472 Toxicrolog
C 6	14.4	90.0	1272	3	Aac50256 Arabidops
C 7	14.4	90.0	1274	3	Aac39069 Arabidops
C 8	14.4	90.0	2169	12	Adq28938 Human CYP
C 9	14.4	90.0	2169	12	Adq38182 Human L-P
C 10	14.4	90.0	2400	10	Abz81308 Human dru
C 11	14.4	90.0	3708	13	Acn42336 Human dia
C 12	14.4	90.0	3779	6	Abn95647 Gene #214
C 13	14.4	90.0	6846	12	Adp68991 Mouse ZAQ
C 14	14.4	90.0	68820	12	Adq97931 Mouse can
C 15	14	87.5	68497	11	Acn45212 Mouse gen
C 16	14	87.5	168325	11	Acn44484 Mouse gen
C 17	13.4	83.8	363	4	Aas32996 DNA encod
C 18	13.4	83.8	400	6	Abn76875 Human ORF
C 19	13.4	83.8	470	3	Aac28376 Human sec
C 20	13.4	83.8	477	4	Aas32997 DNA encod

C 21	13.4	83.8	525	13	ADS51330 Bacterial
C 22	13.4	83.8	535	13	Adr63428 Cotton CD
C 23	13.4	83.8	609	4	Aah90742 CFE 46 co
C 24	13.4	83.8	667	13	Adq55780 Novel can
C 25	13.4	83.8	669	13	Adr91696 Novel S.
C 26	13.4	83.8	712	12	Adl35774 Human PRP
C 27	13.4	83.8	755	13	AdS61997 Bacterial
C 28	13.4	83.8	825	8	ACA50827 Prokaryot
C 29	13.4	83.8	831	2	Adr01681 A. gossyp
C 30	13.4	83.8	1030	9	ACC43774 Nucleotid
C 31	13.4	83.8	1030	10	ABT14524 Escherich
C 32	13.4	83.8	1447	4	AAS59787 Propionib
C 33	13.4	83.8	1447	8	ACF64716 Propionib
C 34	13.4	83.8	1704	10	ADF02304 Bacterial
C 35	13.4	83.8	1712	3	AAS56215 Neospora
C 36	13.4	83.8	1995	10	ACF69455 Photorhab
C 37	13.4	83.8	2289	6	ABS54243 A. sojae
C 38	13.4	83.8	2445	3	Aaz95746 Murine bo
C 39	13.4	83.8	2447	2	Aaq74084 Murine BM
C 40	13.4	83.8	2448	2	AAQ35243 Encodes m
C 41	13.4	83.8	3258	6	ABQ92501 C. pneumo
C 42	13.4	83.8	3258	8	ACA30911 Prokaryot
C 43	13.4	83.8	7415	2	AAX20504 Polynucle
C 44	13.4	83.8	14491	8	ABZ09914 Human 5'
C 45	13.4	83.8	14491	13	ADS89066 Human SDC

## ALIGNMENTS

RESULT 1  
ADJ39000/c  
ID ADJ39000 standard; DNA; 29751 BP.  
XX  
AC ADJ39000;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE SARS coronavirus nucleotide sequence.  
XX  
KW small interfering RNA; siRNA; modified ribonucleotide;  
KW viral replication inhibition; hepatitis C virus; HCV; hepatitis C;  
KW antiinflammatory; hepatotropic; virucide; hepatitis A virus;  
KW hepatitis D virus; hepatitis E virus; Ebola virus; influenza virus;  
KW rotavirus; reovirus; retrovirus; poliovirus; human papilloma virus;  
KW metapneumonavius; coronavirus; viral infection; gene; ds.  
XX- SARS coronavirus.  
OS  
XX WO2004011647-A1.  
XX  
PD 05-FEB-2004.  
XX  
PF 25-JUL-2003; 2003WO-US023104.  
XX  
PR 26-JUL-2002; 2002US-0398605P.  
XX  
PA (CHIR ) CHIRON CORP.  
XX  
PI Han J, Seo MY, Houghton M;  
XX WPI; 2004-143862/14.  
XX  
PT New RNase resistant small interfering RNA, useful for treating viral  
XX infections, e.g., hepatitis C, influenza virus or coronavirus infection.  
XX  
PS Example 10; Fig 3; 74pp; English.  
XX  
CC The present invention describes a small interfering RNA (siRNA) which  
CC comprises a modified ribonucleotide, where the siRNA is resistant to  
CC RNase and retains the ability to inhibit viral replication. Also  
CC described: (1) inactivating a virus in a patient; (2) making a modified  
CC siRNA that targets a nucleic acid sequence in a virus; (3) a double-

CC stranded RNA molecule of 10-30 nucleotides that inhibits replication of  
 CC hepatitis C virus (HCV); (4) inducing targeted RNA interference toward  
 CC HCV in hepatic cells; (5) inhibiting replication of HCV; (6) a vector  
 CC comprising a DNA segment encoding the RNA molecule; (7) a host cell  
 CC comprising the vector of (6); (8) inhibiting replication of HCV in cells  
 CC carrying HCV; (9) treating hepatitis C in a subject; (10) a modified  
 CC siRNA molecule comprising a double-stranded RNA molecule of 10-30  
 CC nucleotides in length, which mediates RNA interference toward a target  
 CC agent or virus and is linked to at least one receptor-binding ligand; and  
 CC (11) inducing targeted RNA interference in a patient. The modified siRNA  
 CC molecules have antiinflammatory, hepatotropic and virucide activities.  
 CC The modified RNA molecules are useful for inactivating virus in mammalian  
 CC cells. The siRNAs are useful for treating hepatitis C virus, hepatitis A  
 CC virus, hepatitis D virus, hepatitis E virus, Ebola virus, influenza  
 CC virus, rotavirus, reovirus, rotovirus, poliovirus, human papilloma  
 CC virus, metapneumoniavirus or coronavirus infections. The methods of the  
 CC invention can be used to correct or compensate for cellular physiological  
 CC abnormalities involved in conferring susceptibility to viral infections  
 CC in patients and/or alleviate symptoms of a viral infection in patients.  
 CC The present sequence represents the SARS coronavirus nucleotide sequence,  
 CC which is used in an example from the present invention.

SQ Sequence 29751 BP; 8478 A; 5941 C; 6188 G; 9144 T; 0 U; 0 Other;

Query Match 100.0%; Score 16; DB 12; Length 29751;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16

Db 17760 TCTGCGTAGGCAATCC 17745

# RESULT 2

ID ABZ51221/c  
 XX ABZ51221 standard; cDNA; 875 BP.

AC ABZ51221;

DT 28-MAR-2003 (first entry)

DE Aspergillus oryzae polynucleotide SEQ ID NO 334.

XX Aspergillus oryzae; fermentation; fungus; industrial; EST;

KW expressed sequence tag; gene; ss.

XX Aspergillus oryzae.

OS Aspergillus oryzae.

XX WO200279476-A1.

PI 10-OCT-2002.

XX 22-MAR-2002; 2002WO-1B000890.

PF 30-MAR-2001; 2001JP-00098371.

XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

PA (NARE-) NAT RES INST BREWING.

PA (NORQ) NAT FOOD RES INST MIN AGRIC.

XX Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;

PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;

XX WPI; 2003-046817/04.

XX Detection of expression of specific Aspergillus genes for monitoring the

PT fermentation and growth conditions of the fungus, using DNA probes.

XX Claim 1; SEQ ID NO 334; 48pp + Sequence Listing; Japanese.

XX The invention relates to a polynucleotide having any of 6006 specific

CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under

CC specific culture conditions including one or more of eutrophic,

CC oligotrophic, solid, early germination, alkaline, high temperature, low  
 CC temperature or maltose culture or polynucleotides stringently hybridising  
 CC to these sequences. The polynucleotides are useful for monitoring the  
 CC progress of fermentation and the growth conditions of a fungus,  
 CC especially of Aspergillus oryzae which is widely used in industrial  
 CC fermentation. Also monitoring for fungal contamination. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 875 BP; 223 A; 219 C; 227 G; 204 T; 0 U; 2 Other;

Query Match 93.8%; Score 15; DB 8; Length 875;

Best Local Similarity 100.0%; Pred. No. 88;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATC 15

Db 661 TCTGCGTAGGCAATC 647

# RESULT 3

AAV31202/c

ID AAV31202 standard; DNA; 746 BP.

XX AAV31202;

DT 01-OCT-1998 (first entry)

XX E. coli J96 pathogenicity island contig #16.

XX PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;

KW PAI V; pheV; vaccine; protective immune response; ds.

XX Escherichia coli.

XX WO9822575-A2.

XX 28-MAY-1998.

XX 21-NOV-1997; 97WO-US021347.

XX 22-NOV-1996; 96US-0031626P.

XX 14-OCT-1997; 97US-0061953P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (UYWI-) UNIV WISCONSIN.

XX Dillon PJ, Choi GH, Welch RA;

XX WPI; 1998-312461/27.

XX New isolated uropathogenic E. coli nucleotide sequences - used to develop

PT products for the detection of pathogenic E. coli and to elicit an immune

PT response to pathogenic E. coli.

XX Claim 21; Page 117; 250pp; English.

XX This sequence represents a E. coli strain J96 contig containing

CC pathogenicity island (PAI) sequences, and represents a nucleic acid

CC molecule of the invention. PAIs are large fragments of DNA which comprise

CC pathogenicity determinants. The sequences of the invention are taken from

CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)

CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at

CC approximately 94 min (at pheR) on the E. coli chromosome and is

CC approximately 160 kb in size. Antibodies specific to the proteins encoded

CC by the PAI open reading frames of the invention can be used in kits to

CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit

CC a protective immune response in an animal to the uropathogenic E. coli

CC strain J96

SQ Sequence 746 BP; 203 A; 150 C; 188 G; 201 T; 0 U; 4 Other;

Query Match 90.0%; Score 14.4; DB 2; Length 746;  
 Best Local Similarity 93.8%; Pred. No. 2e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16  
 |||||  
 Db 334 TCTGCGTAGGCAATCC 319

## RESULT 4

ABZ13359/c  
 ID ABZ13359 standard; DNA; 801 BP.

XX AC ABZ13359;

XX DT 21-JAN-2003 (first entry)

XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1164.

XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX OS Arabidopsis thaliana.

XX PN WO200216655-A2.

XX PD 28-FEB-2002.

XX PF 24-AUG-2001; 2001WO-US026685.

XX PR 24-AUG-2000; 2000US-0227866P.

XX PR 26-JAN-2001; 2001US-0264647P.

XX PR 22-JUN-2001; 2001US-0300111P.

XX PA (SCRI ) SCRIPPS RES INST.

XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX PI Harper JF, Kreps J, Wang X, Zhu T;

XX DR WPI; 2002-304127/34.

XX PT Identifying a stress condition to which a plant cell has been exposed and  
 PT producing plants with increased tolerance to these abiotic stresses.

XX PS Claim 144; SEQ ID NO 1164; 577pp + Sequence Listing; English.

XX CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office

XX SQ Sequence 801 BP; 248 A; 162 C; 178 G; 213 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 6; Length 801;  
 Best Local Similarity 93.8%; Pred. No. 2e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16  
 |||||  
 Db 562 TATGCGTAGGCAATCC 547

## RESULT 5

ABZ83472/c

ID ABZ83472 standard; cDNA; 1024 BP.

XX

AC ABZ83472;

XX DT 14-MAY-2003 (first entry)

XX DE Toxicologically relevant human nucleotide sequence #631.

XX KW Toxicologically relevant gene; toxicological response; gene; ss.

XX OS Homo sapiens.

XX PN WO2003016500-A2.

XX PD 27-FEB-2003.

XX PF 16-AUG-2002; 2002WO-US026514.

XX PR 16-AUG-2001; 2001US-0313080P.

XX PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.

XX PI Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schmeiser K;  
 PI Allen P;

XX DR WPI; 2003-268322/26.

XX CC Determining a toxicological response to an agent, useful for screening of  
 CC drugs, comprises comparing the expression profile of one or more human  
 CC toxic response genes to a reference gene expression profile indicative of  
 CC toxicity.

XX PS Claim 1; Page 215; 455pp; English.

XX CC The present invention describes a method (M1) for determining a  
 CC toxicological response to an agent, which comprises comparing the  
 CC expression profile of one or more human toxic response genes to a  
 CC reference gene expression profile indicative of toxicity, and so  
 CC determining the presence of a toxic response to the agent. Also  
 CC described: (1) an array comprising one or more polynucleotides selected  
 CC from the genes corresponding to the partial sequences given in ABZ82842  
 CC to ABZ84764, or their fragments of at least 20 nucleotides, or homologues  
 CC; and (2) determining if a gene putatively identified to be a toxic  
 CC response gene plays a role on toxic response pathways by determining the  
 CC expression profile of the gene after exposure of cells or a human subject  
 CC to a known toxic pharmaceutical or industrial agent, comprising: (a)  
 CC exposing cells to an agent; (b) obtaining the test gene expression profile  
 CC for a putatively identified toxic response gene after exposure to a known  
 CC toxic pharmaceutical or industrial agent; and (c) comparing the test  
 CC profile to the expression profile of a gene with a similar function or  
 CC comparing the test profile to the expression profile of that gene after  
 CC exposure to other known toxic compounds. The methods are useful for  
 CC predicting and determining toxicological responses on a cellular, organ  
 CC or system level. The arrays comprising the human genes are useful for  
 CC toxicological screening of drugs, pharmaceutical compounds and chemicals

XX SQ Sequence 1024 BP; 271 A; 240 C; 274 G; 239 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 10; Length 1024;  
 Best Local Similarity 93.8%; Pred. No. 2.1e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16  
 |||||  
 Db 357 TCTGCGTAGGCAATCC 342

## RESULT 6

AAC50256/c

ID AAC50256 standard; DNA; 1272 BP.

XX AC AAC50256;

XX

XX DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 64154.  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 28-APR-1999; 99US-0130891P.  
XX 30-APR-1999; 99US-0131449P.  
XX 30-APR-1999; 99US-0132048P.  
XX 30-APR-1999; 99US-0132407P.  
XX 04-MAY-1999; 99US-0132484P.  
XX 05-MAY-1999; 99US-0132485P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132863P.  
XX 11-MAY-1999; 99US-0134256P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 18-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136392P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
XX 03-JUN-1999; 99US-0137528P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.  
XX 16-JUN-1999; 99US-0139453P.  
XX 17-JUN-1999; 99US-0139454P.  
XX 18-JUN-1999; 99US-0139455P.  
XX 18-JUN-1999; 99US-0139456P.  
XX 18-JUN-1999; 99US-0139457P.  
XX 18-JUN-1999; 99US-0139458P.  
XX 18-JUN-1999; 99US-0139459P.  
XX 18-JUN-1999; 99US-0139460P.  
XX 18-JUN-1999; 99US-0139461P.  
XX 18-JUN-1999; 99US-0139462P.  
XX 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
XX 18-JUN-1999; 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
XX 22-JUN-1999; 99US-0139859P.  
XX 23-JUN-1999; 99US-0140353P.  
XX 23-JUN-1999; 99US-0140354P.  
XX 24-JUN-1999; 99US-0140695P.  
XX 28-JUN-1999; 99US-0140823P.  
XX 29-JUN-1999; 99US-0140991P.  
XX 30-JUN-1999; 99US-0141287P.  
XX 01-JUL-1999; 99US-0141842P.  
XX 01-JUL-1999; 99US-0142154P.  
XX 02-JUL-1999; 99US-0142055P.  
XX 06-JUL-1999; 99US-0142390P.  
XX 08-JUL-1999; 99US-0142803P.  
XX 09-JUL-1999; 99US-0142920P.  
XX 12-JUL-1999; 99US-0142977P.  
XX 13-JUL-1999; 99US-0143542P.  
XX 14-JUL-1999; 99US-0143624P.  
XX 15-JUL-1999; 99US-0144005P.  
XX 16-JUL-1999; 99US-0144085P.  
XX 16-JUL-1999; 99US-0144086P.  
XX 19-JUL-1999; 99US-0144325P.  
XX 19-JUL-1999; 99US-0144331P.  
XX 19-JUL-1999; 99US-0144332P.  
XX 19-JUL-1999; 99US-0144333P.  
XX 19-JUL-1999; 99US-0144334P.  
XX 19-JUL-1999; 99US-0144335P.  
XX 20-JUL-1999; 99US-0144352P.  
XX 20-JUL-1999; 99US-0144632P.  
XX 21-JUL-1999; 99US-0144814P.  
XX 21-JUL-1999; 99US-0145086P.  
XX 21-JUL-1999; 99US-0145088P.  
XX 22-JUL-1999; 99US-0145085P.  
XX 22-JUL-1999; 99US-0145087P.  
XX 22-JUL-1999; 99US-0145089P.  
XX 22-JUL-1999; 99US-0145192P.  
XX 23-JUL-1999; 99US-0145145P.  
XX 23-JUL-1999; 99US-0145218P.  
XX 23-JUL-1999; 99US-0145224P.  
XX 26-JUL-1999; 99US-0145276P.  
XX 27-JUL-1999; 99US-0145913P.  
XX 27-JUL-1999; 99US-0145918P.  
XX 27-JUL-1999; 99US-0145919P.  
XX 28-JUL-1999; 99US-0145951P.  
XX 02-AUG-1999; 99US-0146386P.  
XX 02-AUG-1999; 99US-0146388P.  
XX 02-AUG-1999; 99US-0146389P.  
XX 03-AUG-1999; 99US-0147038P.  
XX 04-AUG-1999; 99US-0147204P.  
XX 04-AUG-1999; 99US-0147302P.  
XX 05-AUG-1999; 99US-0147192P.  
XX 05-AUG-1999; 99US-0147260P.  
XX 06-AUG-1999; 99US-0147303P.  
XX 06-AUG-1999; 99US-0147416P.  
XX 09-AUG-1999; 99US-0147493P.  
XX 09-AUG-1999; 99US-0147935P.  
XX 10-AUG-1999; 99US-0148171P.  
XX 11-AUG-1999; 99US-0148319P.  
XX 12-AUG-1999; 99US-0148341P.  
XX 13-AUG-1999; 99US-0148565P.  
XX 16-AUG-1999; 99US-0149368P.  
XX 17-AUG-1999; 99US-0149175P.  
XX 18-AUG-1999; 99US-0149426P.  
XX 20-AUG-1999; 99US-0149722P.  
XX 20-AUG-1999; 99US-0149723P.  
XX 23-AUG-1999; 99US-0149902P.  
XX 23-AUG-1999; 99US-0149930P.  
XX 25-AUG-1999; 99US-0150566P.  
XX 26-AUG-1999; 99US-0150884P.  
XX 27-AUG-1999; 99US-0151065P.  
XX 27-AUG-1999; 99US-0151066P.

```
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155113P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 90.0%; Score 14.4; DB 3; Length 1272;  
Best Local Similarity 93.8%; Pred. No. 2.le+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 TCTGCGTAGGCAATCC 16
    | |||||
Db 779 TATGCGTAGGCAATCC 764
```

## RESULT 7

AAC39069/c  
ID AAC39069 standard; DNA; 1274 BP.

XX

AC AAC39069;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23253.

XX

KW Hybridization assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148566P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151067P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.

PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0160770P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 90.0%; Score 14.4; DB 3; Length 1274;  
Best Local Similarity 93.8%; Pred. No. 2.1e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGGGTAGGCAATCC 16  
| | | | | | | | | | | | | | | |  
Db 781 TATGCGTAGGCAATCC 766

## RESULT 8

ADQ28938/c  
ID ADQ28938 standard; cDNA; 2169 BP.

XX ADQ28938;  
AC ADQ28938;

XX 07-OCT-2004 (first entry)  
DT

XX Human CYP1A2 coding sequence, SEQ ID 9 #1.  
DE

XX Human; CYP1A2; Cytochrome P450; monooxygenase; enzyme; drug metabolism;  
KW gene; ss.  
XX Homo sapiens.

XX OS  
XX WO2004058957-A2.  
PN

XX 15-JUL-2004.  
PD

XX 19-DEC-2003; 2003WO-IB006261.  
PF

XX



PR 31-DEC-2002; 2002US-0437602P.  
 XX (PHAA ) PHARMACIA & UPJOHN CO.  
 XX Gu Y;  
 XX WPI; 2004-525877/50.  
 XX New isolated canine CYP1A2 polypeptide, useful for hybridization assays  
 PT to detect the capacity of cells to express canine CYP1A2 or for measuring  
 PT levels of canine CYP1A2 expression.  
 XX Claim 3; SEQ ID NO 9; 64pp; English.  
 XX The present invention relates to canine CYP1A2 (Cytochrome P450,  
 CC subfamily I, member A2) protein (ADQ28931) and coding sequence  
 CC (ADQ28930). CYP1A2 is a microsomal cytochrome P450 dependent  
 CC monooxygenase which functions in drug metabolism. The sequences useful in  
 CC hybridization assays to detect the capacity of cells to express canine  
 CC CYP1A2 or to measure levels of canine CYP1A2 expression. The present  
 CC sequence was used in a sequence alignment with the canine CYP1A2 coding  
 CC sequence. Note: This sequence is the SEQ ID 9 shown in the sequence  
 CC listing. This sequence differs from the SEQ ID 9 shown in Fig 2  
 CC (ADQ28944).  
 XX SQ Sequence 2169 BP; 589 A; 487 C; 546 G; 547 T; 0 U; 0 Other;  
 Query Match 90.0%; Score 14.4; DB 12; Length 2169;  
 Best Local Similarity 93.8%; Pred. No. 2.2e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TCTGCGTAGGCAATCC 16  
 DB 350 TCTGCGTGGGCAATCC 335  
 RESULT 9  
 ADQ38182/c  
 ID ADQ38182 standard; cDNA; 2169 BP.  
 XX AC ADQ38182;  
 XX DT 07-OCT-2004 (first entry)  
 XX DE Human L-PBE encoding cDNA SEQ ID NO:11.  
 XX KW peroxisomal enoyl-CoA-hydratase/3-hydroxyacyl-CoA dehydrogenase; L-PBE;  
 KW bifunctional enzyme; enzyme; metabolic response; human; gene;  
 KW chromosome 3; ss.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT CDS 1..2169  
 FT /\*tag= a  
 FT /product= "L-PBE"  
 XX WO2004058938-A2.  
 XX PD 15-JUL-2004.  
 XX PF 19-DEC-2003; 2003WO-1B006280.  
 XX PR 31-DEC-2002; 2002US-0437530P.  
 XX PA (PHAA ) PHARMACIA & UPJOHN CO.  
 XX Gu Y;  
 XX WPI; 2004-543453/52.  
 DR P-PSDB; ADQ38183.  
 DR GENBANK; NM\_001966.  
 XX

PT New isolated canine peroxisomal enoyl-CoA-hydratase/3-hydroxyacyl-CoA  
 PT dehydrogenase bifunctional enzyme (L-PBE), useful for measuring the  
 XX metabolic response to a test agent in a dog.  
 XX Claim 3; SEQ ID NO 11; 69pp; English.  
 XX The present invention describes the canine peroxisomal enoyl-CoA-  
 CC hydratase/3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme (L-PBE).  
 CC Also described: (1) an antibody specific for the canine L-PBE polypeptide  
 CC ; (2) an isolated polynucleotide comprising the 2169 bp sequence of SEQ  
 CC ID NO:1 (S1, ADQ38172), or its fragment comprising at least 12  
 CC consecutive nucleotides of S1 or the non-coding strand complementary to  
 CC it with the provision that the fragment comprises a nucleotide sequence  
 CC that differs from any portion of the sequences of 2169, 2154, or 2166 bp  
 CC (see SEQ ID NO:11, 13 and 15, ADQ38182, ADQ38184 and ADQ38186), and from  
 CC their complementary strands by at least one nucleotide; (3) an array of  
 CC nucleic acid molecules, attached to a solid support, the array comprising  
 CC the polynucleotide; (4) an isolated polynucleotide comprising a  
 CC nucleotide sequence that encodes a polypeptide comprising an amino acid  
 CC sequence that is at least 95% homologous to P1 (see SEQ ID NO:2;  
 CC ADQ38173) and which encodes a polypeptide having L-PBE activity; (5) a  
 CC method for determining the amount of canine L-PBE polynucleotide present  
 CC within a sample derived from a dog; (6) a method for measuring the  
 CC metabolic response to a test agent in a dog; and (7) a method for  
 CC determining the amount of canine L-PBE polypeptide present within a  
 CC sample. The canine L-PBE polypeptide and polynucleotide are useful for  
 CC measuring the metabolic response to a test agent in a dog. The present  
 CC sequence encodes human L-PBE, which is used in an example from the  
 CC present invention. The human L-PBE gene is located on chromosome 3, more  
 CC specifically to region 3q26.3-q28.  
 XX SQ Sequence 2169 BP; 589 A; 487 C; 546 G; 547 T; 0 U; 0 Other;  
 Query Match 90.0%; Score 14.4; DB 12; Length 2169;  
 Best Local Similarity 93.8%; Pred. No. 2.2e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TCTGCGTAGGCAATCC 16  
 DB 350 TCTGCGTGGGCAATCC 335  
 RESULT 10  
 ABZ81308/c  
 ID ABZ81308 standard; cDNA; 2400 BP.  
 XX AC ABZ81308;  
 XX DT 10-MAY-2003 (first entry)  
 XX DE Human drug metabolising enzyme, DME-8, coding sequence, SEQ ID 21.  
 XX KW Human; drug metabolising enzyme; anti-HIV; antiallergic;  
 KW antiinflammatory; antianemic; thrombolytic; antilipemic; antidiarrheic;  
 KW antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;  
 KW cyostatic; hepatotropic; virucide; dermatological; antidiabetic;  
 KW nephrotropic; antigout; neuroprotective; thymimetic; osteopathic;  
 KW antiarthritic; antipsoriatic; uropathic; ophthalmological; antirheumatic;  
 KW haemostatic; gene therapy; cell proliferative disorder; cancer;  
 KW developmental disorder; endocrine disorder; eye disorder;  
 KW metabolic disorder; gastrointestinal disorder; liver disorder;  
 KW autoimmune disorder; inflammatory disorder; DME-8; gene; ss.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT CDS 1..2172  
 FT /\*tag= a  
 FT /product= "DME-8"  
 XX WO2003004608-A2.  
 XX PD 16-JAN-2003.

```
XX PF 05-JUL-2002; 2002WO-US0211105.
XX PR 06-JUL-2001; 2001US-0303745P.
XX PR 13-JUL-2001; 2001US-0305402P.
XX PR 27-JUL-2001; 2001US-0308158P.
XX PR 14-SEP-2001; 2001US-0322127P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Griffin JA, Ramkumar J, Emerling BM, Richardson TW, Li JX;
PI Warren BA, Honchell CD, Baughn MR, Tang YT, Lee EA, Elliott VS;
PI Yue H, Lee S, Swarnakar A, Forsythe TJ, Sanjanwala MM, Yao MG;
PI Zebarjadian Y, Gorvad AE, Becha SD, Burford N;
XX WPI; 2003-221588/21.
DR P-PSDB; ABP59217.
XX PT New drug metabolizing enzymes (DME) useful for diagnosing, treating or
PT preventing diseases or conditions associated with aberrant DME
PT expression, e.g. cancer, AIDS, atherosclerosis, diabetes, glaucoma,
PT hepatitis or osteoporosis.
XX Claim 5; Page 177-178; 181pp; English.
XX The present invention relates to novel human drug metabolising enzymes,
CC DME-1 to DME-13 (ABP59210-ABP59222) and their coding sequences (AB281301-
CC AB281313). The sequences are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of DME,
CC particularly cell proliferative disorders (e.g. arteriosclerosis,
CC atherosclerosis, cirrhosis, paroxysmal nocturnal haemoglobinuria,
CC polycythaemia vera, psoriasis, primary thrombocytopenia or cancer),
CC developmental disorders (e.g. renal tubular acidosis, anaemia or mental
CC retardation), endocrine (e.g. osteoporosis, thrombosis, diabetes), eye
CC disorders (e.g. glaucoma, keratitis), metabolic (e.g. hyperlipidaemia,
CC cystic fibrosis), gastrointestinal disorders (e.g. gastroenteritis,
CC diarrhoea), liver disorders (e.g. hepatitis, Reye's syndrome), or
CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, asthma,
CC autoimmune thyroiditis, contact dermatitis, Crohn's disease,
CC glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease,
CC Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, pancreatitis, Reiter's syndrome, rheumatoid arthritis,
CC Sjogren's syndrome, uveitis). They are also useful in the assessing the
CC effects of exogenous compounds on the expression of nucleic acid and
CC amino acid sequences of DME. The polynucleotides encoding DME are useful
CC for creating transgenic animals to model human disease
XX SQ Sequence 2400 BP; 635 A; 539 C; 603 G; 623 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 10; Length 2400;
Best Local Similarity 93.8%; Pred. No. 2.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
Db 350 TCTGCGTAGGCAATCC 335

RESULT 11
ACN42336/c
ID ACN42336 standard; cDNA; 3708 BP.
XX AC ACN42336;
XX DT 18-NOV-2004 (first entry)
XX DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:1211.
XX ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
XX dithp.
XX KW metastatic liver tumour; cytostatic; expression profile; disease state;
XX disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.

PN WO2004023973-A2.
XX 25-MAR-2004.
XX 12-SEP-2003; 2003WO-US028227.
XX 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Eldred LV;
XX Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX Patuary S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
DR P-PSDB; ABM83684.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX Claim 1; Page; 190pp; English.
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp polynucleotide of
CC the invention. Note: The sequence data for this patent is not represented
CC in the printed specification, but was obtained in electronic format
CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 3708 BP; 1119 A; 724 C; 837 G; 1028 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 13; Length 3708;
Best Local Similarity 93.8%; Pred. No. 2.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
Db 390 TCTGCGTAGGCAATCC 375

RESULT 12
ABN95647/c
ID ABN95647 standard; DNA; 3779 BP.
XX AC ABN95647;
XX DT 13-AUG-2002 (first entry)
XX DE Gene #2145 used to diagnose liver cancer.
XX ss; gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX metastatic liver tumour; cytostatic; expression profile; disease state;
XX disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.
```



PT New isolated cancer associated nucleic acids comprising at least 10  
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating  
XX cancers such as leukemia and lymphoma.

PS Claim 1; SEQ ID NO 908; 199pp; English.

XX The present invention relates to cancer associated sequences (ADQ97025-  
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or  
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 68820 BP; 17966 A; 14439 C; 15095 G; 20194 T; 0 U; 1126 Other;

Query Match 90.0%; Score 14.4; DB 12; Length 68820;

Best Local Similarity 93.8%; Pred. No. 3.2e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16

|||||

Db 20254 TTTGCGTAGGCAATCC 20269

RESULT 15

ACN45212

ID ACN45212 standard; DNA; 68497 BP.

XX ACN45212;

DT 18-NOV-2004 (first entry)

DE Mouse genomic sequence mCG9808.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.

XX Mus musculus.

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
XX comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 2047; Opp; English.

XX The present invention relates to novel DNA and protein sequences which  
CC are associated with carcinomas. The sequences are useful for: (i) for  
CC screening drug candidates; (ii) for screening of bioactive agent capable  
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
CC carcinoma including lymphoma. The present sequence is one such CA coding  
CC sequence. Note: This patent is an equivalent to basic patent  
CC US2002182586A1, for which no sequence data was published

XX Sequence 68497 BP; 16500 A; 15015 C; 15455 G; 18939 T; 0 U; 2588 Other;

Query Match 87.5%; Score 14; DB 11; Length 68497;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATC 15

|||||

Db 43582 CTGCGTAGGCAATC 43595

Search completed: May 16, 2005, 01:09:26  
Job time : 210.25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:11:05 ; Search time 59.875 Seconds  
(without alignments)  
437.251 Million cell updates/sec

Title: US-10-808-187A-2473

Perfect score: 16

Sequence: 1 tctgtaggcaatcc 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14.4	90.0	746	3	US-08-976-259-16
C 2	14.4	90.0	746	4	US-09-956-004-16
C 3	13.4	83.8	470	4	US-09-513-999C-32451
C 4	13.4	83.8	601	4	US-09-949-016-113217
C 5	13.4	83.8	669	4	US-09-107-433-331
C 6	13.4	83.8	831	3	US-08-998-416-373
C 7	13.4	83.8	1704	4	US-09-543-681A-2589
C 8	13.4	83.8	2200	3	US-09-504-358-27
C 9	13.4	83.8	2200	3	US-09-954-314-27
C 10	13.4	83.8	2200	4	US-10-230-562-27
C 11	13.4	83.8	2447	1	US-08-050-132A-1
C 12	13.4	83.8	2447	3	US-08-750-222A-1
C 13	13.4	83.8	2447	3	US-08-815-652B-1
C 14	13.4	83.8	2447	3	US-08-254-353A-1
C 15	13.4	83.8	2447	5	PCT-US92-05374A-1
C 16	13.4	83.8	2447	5	PCT-US95-07084-1
C 17	13.4	83.8	11471	3	US-09-504-358-16
C 18	13.4	83.8	11471	3	US-09-954-314-16
C 19	13.4	83.8	11471	4	US-10-230-562-16
C 20	13.4	83.8	27120	4	US-09-949-016-16210
C 21	13.4	83.8	62804	3	US-09-800-960-3
C 22	13.4	83.8	62804	4	US-10-096-960-3
C 23	13.4	83.8	66933	4	US-09-544-398B-11
C 24	13.4	83.8	66933	4	US-09-543-771B-11
C 25	13.4	83.8	72049	4	US-09-544-398B-9
C 26	13.4	83.8	72049	4	US-09-543-771B-9
C 27	13.4	83.8	140844	4	US-09-949-016-14199

C 28	13.4	83.8	152524	4	US-09-949-016-12683	Sequence 12683, A
C 29	13.4	83.8	152524	4	US-09-949-016-13194	Sequence 13194, A
C 30	13.4	83.8	285986	4	US-09-949-016-12287	Sequence 12287, A
C 31	13.4	83.8	288031	4	US-09-949-016-14864	Sequence 14864, A
C 32	13.4	83.8	317366	4	US-09-949-016-16001	Sequence 16001, A
C 33	13.4	83.8	1230025	4	US-09-198-452A-1	Sequence 1, Appli
C 34	13.4	83.8	1230230	4	US-09-438-185A-1	Sequence 1, Appli
C 35	13.4	83.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 36	13.4	83.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C 37	13	81.2	870	4	US-09-543-681A-2890	Sequence 2890, A
C 38	13	81.2	1006	4	US-09-603-208A-183	Sequence 183, App
C 39	13	81.2	1341	4	US-09-603-208A-181	Sequence 181, App
C 40	13	81.2	1368	4	US-09-602-787A-529	Sequence 529, App
C 41	13	81.2	2715	4	US-09-248-796A-1444	Sequence 1444, Ap
C 42	13	81.2	5240	3	US-09-171-337A-2	Sequence 2, Appli
C 43	13	81.2	5240	4	US-09-631-022-2	Sequence 2, Appli
C 44	12.8	80.0	25	4	US-09-396-196G-5481	Sequence 5481, Ap
C 45	12.8	80.0	25	4	US-09-396-196G-5482	Sequence 5482, Ap

ALIGNMENTS

RESULT 1

US-08-976-259-16/c

; Sequence 16, Application US/08976259

; Patent No. 6316609

; GENERAL INFORMATION:

; APPLICANT: Dillion, Patrick J.

; APPLICANT: Choi, Gil H.

; APPLICANT: Welch, Rodney A.

; TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli

; NUMBER OF SEQUENCES: 142

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

; STREET: 1100 New York Ave, N.W., Suite 600

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/976,259

; FILING DATE: Herewith

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/031,626 AND US 60/061,953

; ATTORNEY/AGENT INFORMATION:

; NAME: Steffe, Eric K.

; REGISTRATION NUMBER: 36,688

; REFERENCE/DOCKET NUMBER: 1488.0740002/EKS/CBM

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 746 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-976-259-16

Query Match 90.0%; Score 14.4; DB 3; Length 746;

Best Local Similarity 93.8%; Pred. No. 70;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCTAGGCAATCC 16

|||||||

Db 334 TCTGCGTAGCCAATCC 319

## RESULT 2

US-09-956-004-16/c  
; Sequence 16, Application US/09956004  
; Patent No. 6787643  
; GENERAL INFORMATION:  
; APPLICANT: Patrick J. Dillon et al.  
; TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands  
; FILE REFERENCE: PB324D1  
; CURRENT APPLICATION NUMBER: US/09/956,004  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: 08/976,259  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/061,953  
; PRIOR FILING DATE: 1997-10-14  
; PRIOR APPLICATION NUMBER: 60/031,626  
; PRIOR FILING DATE: 1996-11-22  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 746  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (718)..(718)  
; OTHER INFORMATION: n equals a, t, g, or c  
; NAME/KEY: misc feature  
; LOCATION: (741)..(741)  
; OTHER INFORMATION: n equals a, t, g, or c  
US-09-956-004-16

Query Match 90.0%; Score 14.4; DB 4; Length 746;

Best Local Similarity 93.8%; Pred. No. 70;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCGTAGCCAATCC 16

Db 334 TCTGCGTAGCCAATCC 319

## RESULT 3

US-09-513-999C-32451/c  
; Sequence 32451, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 32451  
; LENGTH: 470  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-513-999C-32451

Query Match 83.8%; Score 13.4; DB 4; Length 470;

Best Local Similarity 93.3%; Pred. No. 2.6e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCGTAGCCAATC 15

Db 317 TCTGCGTAGCCACTC 303

## RESULT 4

US-09-949-016-113217/c  
; Sequence 113217, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 113217  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human'  
US-09-949-016-113217

Query Match 83.8%; Score 13.4; DB 4; Length 601;

Best Local Similarity 93.3%; Pred. No. 2.7e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGCGTAGCCAATC 15

Db 149 TCTGCATAGGCAATC 135

## RESULT 5

US-09-107-433-331/c  
; Sequence 331, Application US/09107433  
; Patent No. 6800744  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID  
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE  
; THERAPEUTICS  
; NUMBER OF SEQUENCES: 5206  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: <Unknown>  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: <Unknown>  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,433  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/ 085131  
; FILING DATE: May 12, 1998  
; APPLICATION NUMBER: 60/051553  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277

```
; INFORMATION FOR SEQ ID NO: 331:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 669 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;     NAME/KEY: misc feature
;     LOCATION: (B) LOCATION 1...669
; SEQUENCE DESCRIPTION: SEQ ID NO: 331:
US-09-107-433-331

Query Match      83.8%; Score 13.4; DB 4; Length 669;
Best Local Similarity 93.3%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CTGCGTAGGCAATCC 16
      ||||| |||||
Db      165 CTGCGTCGCAATCC 151

RESULT 6
US-08-998-416-373/c
; Sequence 373, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
;   APPLICANT: Philippsen, Peter
;   APPLICANT: Pohlmann, Rainer
;   APPLICANT: Steinert, Sabine
;   APPLICANT: Mohr, Christine
;   APPLICANT: Wendland, Jurgen
;   APPLICANT: Knechtie, Philipp
;   APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: No. 6239264artis Corporation
;   STREET: 3054 Cornwallis Road
;   CITY: Research Triangle Park
;   STATE: No. 6239264th Carolina
;   COUNTRY: USA
;   ZIP: 27709
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/998,416
;   FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: CH 0016/97
;   FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME: Meigs, J. Timothy
;   REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 919-541-8587
;   TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 373:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 831 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
```

```
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
;   ORGANISM: PAG1284UP
; US-08-998-416-373

Query Match      83.8%; Score 13.4; DB 3; Length 831;
Best Local Similarity 93.3%; Pred. No. 2.8e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CTGCGTAGGCAATCC 16
      ||||| |||||
Db      425 CTGCGTTGGCAATCC 411

RESULT 7
US-09-543-681A-2589
; Sequence 2589, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;   APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2589
; LENGTH: 1704
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-2589

Query Match      83.8%; Score 13.4; DB 4; Length 1704;
Best Local Similarity 93.3%; Pred. No. 3.2e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CTGCGTAGGCAATCC 16
      ||||| |||||
Db      706 CTGCGTTGGCAATCC 720

RESULT 8
US-09-504-358-27
; Sequence 27, Application US/09504358
; Patent No. 6365376
; GENERAL INFORMATION:
;   APPLICANT: Rouviere, Pierre E.
;   APPLICANT: Brzostowicz, Patricia C.
; TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIATE
; FILE REFERENCE: BC1001 US NA
; CURRENT APPLICATION NUMBER: US/09/504,358
; CURRENT FILING DATE: 2000-02-15
; EARLIER APPLICATION NUMBER: 60/120,702
; EARLIER FILING DATE: 1999-February-19
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Brevibacterium sp HCU
; US-09-504-358-27

Query Match      83.8%; Score 13.4; DB 3; Length 2200;
Best Local Similarity 93.3%; Pred. No. 3.3e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CTGCGTAGGCAATCC 16
      ||||| |||||
Db      644 CCGCGTAGGCAATCC 658
```

```
RESULT 9
US-09-954-314-27
; Sequence 27, Application US/09954314
; Patent No. 6465224
; GENERAL INFORMATION:
; APPLICANT: Rouviere, Pierre E.
; APPLICANT: Brzostowicz, Patricia C.
; TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIATES
; FILE REFERENCE: BC1001 US NA
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/120,702
; PRIOR FILING DATE: 1999-February-19
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; TYPE: DNA
; ORGANISM: Brevibacterium sp HCU
US-09-954-314-27

Query Match      83.8%; Score 13.4; DB 3; Length 2200;
Best Local Similarity 93.3%; Pred. No. 3.3e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CTGCGTAGGCAATCC 16
      |||||
DB      644 CCGCGTAGGCAATCC 658

RESULT 10
US-10-230-562-27
; Sequence 27, Application US/10230562
; Patent No. 6790645
; GENERAL INFORMATION:
; APPLICANT: Rouviere, Pierre E
; APPLICANT: Brzostowicz, Patricia C
; TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID
; FILE REFERENCE: BC-1001
; CURRENT APPLICATION NUMBER: US/10/230,562
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/120,702
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Brevibacterium sp HCU
US-10-230-562-27

Query Match      83.8%; Score 13.4; DB 4; Length 2200;
Best Local Similarity 93.3%; Pred. No. 3.3e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CTGCGTAGGCAATCC 16
      |||||
DB      644 CCGCGTAGGCAATCC 658

RESULT 11
US-08-050-132A-1/c
; Sequence 1, Application US/08050132A
; Patent No. 5661007
; GENERAL INFORMATION:
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony
; TITLE OF INVENTION: BMP-9 COMPOSITIONS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: Legal Affairs - 87 CambridgePark Drive

US-08-050-132A-1/c
; Sequence 1, Application US/08750222A
; Patent No. 6034061
; GENERAL INFORMATION:
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Song, Jeffrey
; APPLICANT: Thies, Scott
; TITLE OF INVENTION: BMP-9 COMPOSITIONS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: Legal Affairs - 87 CambridgePark Drive
```

```
; CITY: Cambridge
; STATE: MA
; COUNTRY: US
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/050,132A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI 5186A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; STRAIN: C57B46xCBA
; TISSUE TYPE: liver
; IMMEDIATE SOURCE:
; LIBRARY: Mouse liver cdna
; CLONE: ML14A
; POSITION IN GENOME:
; UNITS: bp
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1564..1893
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 610..1896
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: 1..2447
; US-08-050-132A-1

Query Match      83.8%; Score 13.4; DB 1; Length 2447;
Best Local Similarity 93.3%; Pred. No. 3.3e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CTGCGTAGGCAATCC 16
      |||||
DB      812 CTGCGTAGGCAATCC 798

RESULT 12
US-08-750-222A-1/c
; Sequence 1, Application US/08750222A
; Patent No. 6034061
; GENERAL INFORMATION:
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Song, Jeffrey
; APPLICANT: Thies, Scott
; TITLE OF INVENTION: BMP-9 COMPOSITIONS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: Legal Affairs - 87 CambridgePark Drive
```



;; CITY: Cambridge  
;; STATE: MA  
;; COUNTRY: US  
;; ZIP: 02140  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; FILING DATE: 04-DEC-1996  
;; APPLICATION NUMBER: US/08/750,222A  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/254,353  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kapinos, Ellen J.  
;; REGISTRATION NUMBER: 32,245  
;; REFERENCE/DOCKET NUMBER: GI 5186B  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 876-1170  
;; TELEFAX: (617) 876-5851  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2447 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA to mRNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: Mus musculus  
;; STRAIN: C57B46xCBA  
;; TISSUE TYPE: liver  
;; IMMEDIATE SOURCE:  
;; LIBRARY: Mouse liver cDNA  
;; CLONE: ML14A  
;; POSITION IN GENOME:  
;; UNITS: bp  
;; FEATURE:  
;; NAME/KEY: mat\_peptide  
;; LOCATION: 1564..1893  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 610..1896  
;; FEATURE:  
;; NAME/KEY: mRNA  
;; LOCATION: 1..2447  
US-08-750-222A-1

Query Match 83.8%; Score 13.4; DB 3; Length 2447;  
Best Local Similarity 93.3%; Pred. No. 3.3e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16  
|||||  
Db 812 CTGCGTAGGCAATCC 798

RESULT 13  
US-08-815-652B-1/c  
; Sequence 1, Application US/08815652B  
; Patent No. 6034062  
; GENERAL INFORMATION:  
; APPLICANT: Wozney, John M.  
; APPLICANT: Celeste, Anthony  
; APPLICANT: Song, Jeffrey  
; APPLICANT: Thies, R. Scott  
; TITLE OF INVENTION: BMP-9 COMPOSITIONS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Genetics Institute, Inc.  
;; STREET: Legal Affairs - 87 CambridgePark Drive  
;; CITY: Cambridge  
;; STATE: MA  
;; COUNTRY: US  
;; ZIP: 02140  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/815,652B  
;; FILING DATE:  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kapinos, Ellen J.  
;; REGISTRATION NUMBER: 32,245  
;; REFERENCE/DOCKET NUMBER: GI 5186D  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 876-1170  
;; TELEFAX: (617) 876-5851  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2447 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA to mRNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: Mus musculus  
;; STRAIN: C57B46xCBA  
;; TISSUE TYPE: liver  
;; IMMEDIATE SOURCE:  
;; LIBRARY: Mouse liver cDNA  
;; CLONE: ML14A  
;; POSITION IN GENOME:  
;; UNITS: bp  
;; FEATURE:  
;; NAME/KEY: mat\_peptide  
;; LOCATION: 1564..1893  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 610..1896  
;; FEATURE:  
;; NAME/KEY: mRNA  
;; LOCATION: 1..2447  
US-08-815-652B-1

Query Match 83.8%; Score 13.4; DB 3; Length 2447;  
Best Local Similarity 93.3%; Pred. No. 3.3e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16  
|||||  
Db 812 CTGCGTAGGCAATCC 798

RESULT 14  
US-08-254-353A-1/c  
; Sequence 1, Application US/08254353A  
; Patent No. 6287816  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Vicki A.  
; APPLICANT: Wozney, John M.  
; APPLICANT: Celeste, Anthony J.  
; APPLICANT: Song, Jeffrey  
; APPLICANT: Thies, Scott  
; TITLE OF INVENTION: BMP-9 COMPOSITIONS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.  
STREET: Legal Affairs - 87 CambridgePark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: US  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/254,353A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kapinos, Ellen J.  
REGISTRATION NUMBER: 32,245  
REFERENCE/DOCKET NUMBER: GI 5186B  
TELEPHONE: (617) 876-1170  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
STRAIN: C57B46xCBA  
TISSUE TYPE: liver  
IMMEDIATE SOURCE:  
LIBRARY: Mouse liver cDNA  
CLONE: ML14A  
POSITION IN GENOME:  
UNITS: bp  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1564..1893  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 610..1896  
FEATURE:  
NAME/KEY: mRNA  
LOCATION: 1..2447  
US-08-254-353A-1

Query Match 83.8%; Score 13.4; DB 3; Length 2447;  
Best Local Similarity 93.3%; Pred. No. 3.3e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16  
Db 812 CTGCGTAGGCAATCC 798

RESULT 15  
PCT-US92-05374A-1/c  
Sequence 1, Application PC/TUS9205374A  
GENERAL INFORMATION:  
APPLICANT: Wozney, John M.  
APPLICANT: Celeste, Anthony  
TITLE OF INVENTION: BMP-9 COMPOSITIONS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: Legal Affairs - 87 CambridgePark Drive  
CITY: Cambridge  
STATE: MA

COUNTRY: US  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/05374A  
FILING DATE: 19920625  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kapinos, Ellen J.  
REGISTRATION NUMBER: 32,245  
REFERENCE/DOCKET NUMBER: GI 5186A  
TELEPHONE: (617) 876-1170  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2447 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
STRAIN: C57B46xCBA  
TISSUE TYPE: liver  
IMMEDIATE SOURCE:  
LIBRARY: Mouse liver cDNA  
CLONE: ML14A  
POSITION IN GENOME:  
UNITS: bp  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1564..1893  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 610..1896  
FEATURE:  
NAME/KEY: mRNA  
LOCATION: 1..2447  
PCT-US92-05374A-1

Query Match 83.8%; Score 13.4; DB 5; Length 2447;  
Best Local Similarity 93.3%; Pred. No. 3.3e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16  
Db 812 CTGCGTAGGCAATCC 798

Search completed: May 16, 2005, 06:11:37  
Job time : 66.875 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 02:00:36 ; Search time 758.5 Seconds  
(without alignments)  
129.102 Million cell updates/sec

Title: US-10-808-187A-2473

Perfect score: 16

Sequence: 1 tctgcgtaggcaatcc 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : \* Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	100.0	16	19	US-10-808-187-2473
C 2	16	100.0	646	19	Sequence 2473, Ap
C 3	16	100.0	1213	19	Sequence 1, Appli
C 4	16	100.0	5262	19	Sequence 11, Appl
C 5	16	100.0	28920	19	Sequence 9, Appli
C 6	16	100.0	28920	19	Sequence 5, Appli
C 7	16	100.0	28920	19	Sequence 6, Appli
C 8	16	100.0	29291	19	Sequence 4, Appli
C 9	16	100.0	29430	19	Sequence 7, Appli
C 10	16	100.0	29727	18	US-10-839-729-15
C 11	16	100.0	29727	18	Sequence 15, Appl
					Sequence 1, Appli
					Sequence 8, Appli

C 12	16	100.0	29727	19	US-10-699-936-1
C 13	16	100.0	29736	18	US-10-839-729-17
C 14	16	100.0	29736	19	US-10-889-447-9
C 15	16	100.0	29736	19	US-10-699-936-3
C 16	16	100.0	29742	18	US-10-839-729-16
C 17	16	100.0	29742	19	US-10-808-187-15
C 18	16	100.0	29742	19	US-10-808-187-16
C 19	16	100.0	29742	19	US-10-808-187-240
C 20	16	100.0	29742	19	US-10-808-187-737
C 21	16	100.0	29742	19	US-10-808-187-1108
C 22	16	100.0	29742	19	US-10-808-187-1590
C 23	16	100.0	29742	19	US-10-808-187-1965
C 24	16	100.0	29742	19	US-10-889-447-10
C 25	16	100.0	29751	18	US-10-839-729-14
C 26	16	100.0	29751	19	US-10-856-529-1
C 27	16	100.0	29751	19	US-10-626-879-67
C 28	16	100.0	29751	19	US-10-889-447-1
C 29	16	100.0	29751	19	US-10-889-447-2
C 30	16	100.0	29751	19	US-10-699-936-2
C 31	14.4	90.0	746	9	US-09-956-004-16
C 32	14.4	90.0	746	18	US-10-808-570-16
C 33	14.4	90.0	801	9	US-09-938-842A-1164
C 34	14.4	90.0	801	11	US-09-938-842A-1164
C 35	14.4	90.0	1032	18	US-10-437-963-27078
C 36	14.4	90.0	2169	18	US-10-743-884-11
C 37	14.4	90.0	2169	18	US-10-744-150-9
C 38	14.4	90.0	3779	9	US-09-880-107-2145
C 39	14	87.5	491	18	US-10-425-115-25291
C 40	14	87.5	2547	18	US-10-425-115-87839
C 41	14	87.5	2898	18	US-10-425-115-87795
C 42	14	87.5	3474	18	US-10-425-115-87943
C 43	14	87.5	3609	18	US-10-425-115-87877
C 44	14	87.5	3909	18	US-10-425-115-87914
C 45	14	87.5	3994	18	US-10-425-115-87810

#### ALIGNMENTS

#### RESULT 1

US-10-808-187-2473  
; Sequence 2473, Application US/10808187  
; Publication No. US2005009009A1  
; GENERAL INFORMATION:  
; APPLICANT: PEIRIS, JOSEPH S. M.  
; APPLICANT: YUEN, KWOK YUNG  
; APPLICANT: POON, LIT MAN  
; APPLICANT: GUAN, YI  
; APPLICANT: CHAN, KWOK HUNG  
; TITLE OF INVENTION: NICHOLLS, JOHN  
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE  
; TITLE OF INVENTION: RESPIRATORY SYNDROME (SARS)  
; FILE REFERENCE: V9661.0078  
; CURRENT APPLICATION NUMBER: US/10/808,187  
; CURRENT FILING DATE: 2004-03-24  
; PRIOR APPLICATION NUMBER: 60/457,031  
; PRIOR FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: 60/457,730  
; PRIOR FILING DATE: 2003-03-26  
; PRIOR APPLICATION NUMBER: 60/459,931  
; PRIOR FILING DATE: 2003-04-02  
; PRIOR APPLICATION NUMBER: 60/460,357  
; PRIOR FILING DATE: 2003-04-03  
; PRIOR APPLICATION NUMBER: 60/461,265  
; PRIOR FILING DATE: 2003-04-08  
; PRIOR APPLICATION NUMBER: 60/462,805  
; PRIOR FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 60/468,139  
; PRIOR FILING DATE: 2003-05-05  
; PRIOR APPLICATION NUMBER: 60/464,886  
; PRIOR FILING DATE: 2003-04-23  
; PRIOR APPLICATION NUMBER: 60/471,200  
; PRIOR FILING DATE: 2003-05-16

; NUMBER OF SEQ ID NOS: 2476  
; SOFTWARE: PatentIn ver. 3.2  
; SEQ ID NO 2473  
; LENGTH: 16  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Probe  
US-10-808-187-2473

Query Match 100.0%; Score 16; DB 19; Length 16;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16  
|||||  
Db 1 TCTGCGTAGGCAATCC 16

RESULT 2  
US-10-808-187-1/c  
; Sequence 1, Application US/10808187  
; Publication No. US200500909A1  
; GENERAL INFORMATION:  
; APPLICANT: PEIRIS, JOSEPH S. M.  
; APPLICANT: YUEN, KWOK YUNG  
; APPLICANT: POON, LIT MAN  
; APPLICANT: GUAN, YI  
; APPLICANT: CHAN, KWOK HUNG  
; APPLICANT: NICHOLLS, JOHN  
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE  
; FILE REFERENCE: V9661.0078  
; CURRENT APPLICATION NUMBER: US/10/808,187  
; PRIOR FILING DATE: 2004-03-24  
; PRIOR APPLICATION NUMBER: 60/457,031  
; PRIOR FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: 60/457,730  
; PRIOR FILING DATE: 2003-03-26  
; PRIOR APPLICATION NUMBER: 60/459,931  
; PRIOR FILING DATE: 2003-04-02  
; PRIOR APPLICATION NUMBER: 60/460,357  
; PRIOR FILING DATE: 2003-04-03  
; PRIOR APPLICATION NUMBER: 60/461,265  
; PRIOR FILING DATE: 2003-04-08  
; PRIOR APPLICATION NUMBER: 60/462,805  
; PRIOR FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 60/468,139  
; PRIOR FILING DATE: 2003-05-05  
; PRIOR APPLICATION NUMBER: 60/464,886  
; PRIOR FILING DATE: 2003-04-23  
; PRIOR APPLICATION NUMBER: 60/471,200  
; PRIOR FILING DATE: 2003-05-16  
; NUMBER OF SEQ ID NOS: 2476  
; SOFTWARE: PatentIn ver. 3.2  
; SEQ ID NO 1  
; LENGTH: 646  
; TYPE: DNA  
; ORGANISM: Human severe acute respiratory system virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)...(646)  
US-10-808-187-1

Query Match 100.0%; Score 16; DB 19; Length 646;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16  
|||||  
Db 44 TCTGCGTAGGCAATCC 29

RESULT 3  
US-10-808-187-11/c  
; Sequence 11, Application US/10808187  
; Publication No. US200500909A1  
; GENERAL INFORMATION:  
; APPLICANT: PEIRIS, JOSEPH S. M.  
; APPLICANT: YUEN, KWOK YUNG  
; APPLICANT: POON, LIT MAN  
; APPLICANT: GUAN, YI  
; APPLICANT: CHAN, KWOK HUNG  
; APPLICANT: NICHOLLS, JOHN  
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE  
; FILE REFERENCE: V9661.0078  
; CURRENT APPLICATION NUMBER: US/10/808,187  
; CURRENT FILING DATE: 2004-03-24  
; PRIOR APPLICATION NUMBER: 60/457,031  
; PRIOR FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: 60/457,730  
; PRIOR FILING DATE: 2003-03-26  
; PRIOR APPLICATION NUMBER: 60/459,931  
; PRIOR FILING DATE: 2003-04-02  
; PRIOR APPLICATION NUMBER: 60/460,357  
; PRIOR FILING DATE: 2003-04-03  
; PRIOR APPLICATION NUMBER: 60/461,265  
; PRIOR FILING DATE: 2003-04-08  
; PRIOR APPLICATION NUMBER: 60/462,805  
; PRIOR FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 60/468,139  
; PRIOR FILING DATE: 2003-05-05  
; PRIOR APPLICATION NUMBER: 60/464,886  
; PRIOR FILING DATE: 2003-04-23  
; PRIOR APPLICATION NUMBER: 60/471,200  
; PRIOR FILING DATE: 2003-05-16  
; NUMBER OF SEQ ID NOS: 2476  
; SOFTWARE: PatentIn ver. 3.2  
; SEQ ID NO 11  
; LENGTH: 1213  
; TYPE: DNA  
; ORGANISM: Human severe acute respiratory system virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)...(1213)  
US-10-808-187-11

Query Match 100.0%; Score 16; DB 19; Length 1213;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16  
|||||  
Db 611 TCTGCGTAGGCAATCC 596

RESULT 4  
US-10-699-936-9/c  
; Sequence 9, Application US/10699936  
; Publication No. US20050095582A1  
; GENERAL INFORMATION:  
; APPLICANT: Gillim-Ross, Laura  
; APPLICANT: Taylor, Jill  
; APPLICANT: Scholl, David R.  
; APPLICANT: Jollick, Joseph D.  
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory  
; FILE REFERENCE: DHI-07986  
; CURRENT APPLICATION NUMBER: US/10/699,936  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 9  
; LENGTH: 5262

```
; TYPE: DNA
; ORGANISM: SARS coronavirus Shanghai QXC
US-10-699-936-9

Query Match      100.0%; Score 16; DB 19; Length 5262;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
Db 2026 TCTGCGTAGGCAATCC 2011

RESULT 5
US-10-889-447-5/c
; Sequence 5, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ03
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-5

Query Match      100.0%; Score 16; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
Db 17487 TCTGCGTAGGCAATCC 17472

RESULT 6
US-10-889-447-6/c
; Sequence 6, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ04
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-6

Query Match      100.0%; Score 16; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
Db 17487 TCTGCGTAGGCAATCC 17472

RESULT 7
US-10-889-447-4/c
; Sequence 4, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 29291
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ02
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-4

Query Match      100.0%; Score 16; DB 19; Length 29291;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
Db 17813 TCTGCGTAGGCAATCC 17798

RESULT 8
US-10-889-447-7/c
; Sequence 7, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 29430
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate GZ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-7

Query Match      100.0%; Score 16; DB 19; Length 29430;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
Db 17813 TCTGCGTAGGCAATCC 17798
```

```
Db 17472 TCTGCGTAGGCAATCC 17457

RESULT 9
US-10-839-729-15/c
; Sequence 15, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
; FILE REFERENCE: BIOBANK.013A
; CURRENT APPLICATION NUMBER: US/10/839,729
; CURRENT FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: 60/468703
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-839-729-15

Query Match 100.0%; Score 16; DB 18; Length 29727;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
|||||
Db 17760 TCTGCGTAGGCAATCC 17745

RESULT 10
US-10-827-757-1/c
; Sequence 1, Application US/10827757
; Publication No. US20050004071A1
; GENERAL INFORMATION:
; APPLICANT: Comper, Wayne
; TITLE OF INVENTION: Charged Polysaccharides Resistant To Lysosomal Degradation During
; TITLE OF INVENTION: Kidney Filtration And Renal Passage And Their Use To Treat Or
; TITLE OF INVENTION: Prevent Infection By Coronaviruses
; FILE REFERENCE: 11213-007-999
; CURRENT APPLICATION NUMBER: US/10/827,757
; CURRENT FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 60/464,294
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS-related coronavirus (Urbani strain)
US-10-827-757-1

Query Match 100.0%; Score 16; DB 18; Length 29727;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
|||||
Db 17760 TCTGCGTAGGCAATCC 17745

RESULT 11
US-10-889-447-8/c
; Sequence 8, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US

; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS coronavirus Urbani
US-10-889-447-8

Query Match 100.0%; Score 16; DB 19; Length 29727;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
|||||
Db 17760 TCTGCGTAGGCAATCC 17745

RESULT 12
US-10-699-936-1/c
; Sequence 1, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS coronavirus Urbani
US-10-699-936-1

Query Match 100.0%; Score 16; DB 19; Length 29727;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16
|||||
Db 17760 TCTGCGTAGGCAATCC 17745

RESULT 13
US-10-839-729-17/c
; Sequence 17, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
; FILE REFERENCE: BIOBANK.013A
; CURRENT APPLICATION NUMBER: US/10/839,729
; CURRENT FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: 60/468703
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 29736
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-839-729-17
```

Job time : 759.5 secs

Query Match 100.0%; Score 16; DB 18; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCGTAGGCAATCC 16  
Db 17745 TCTGCGTAGGCAATCC 17730

RESULT 14

US-10-889-447-9/c  
; Sequence 9, Application US/10889447  
; Publication No. US20050075307A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Jain, Ravi  
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
; FILE REFERENCE: RTS-0685US  
; CURRENT APPLICATION NUMBER: US/10/889,447  
; PRIOR FILING DATE: 2004-07-12  
; PRIOR APPLICATION NUMBER: 60/486,670  
; PRIOR FILING DATE: 2003-07-12  
; NUMBER OF SEQ ID NOS: 241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 29736  
; TYPE: DNA  
; ORGANISM: SARS coronavirus CUHK-W1  
US-10-889-447-9

Query Match 100.0%; Score 16; DB 19; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCGTAGGCAATCC 16  
Db 17745 TCTGCGTAGGCAATCC 17730

RESULT 15

US-10-699-936-3/c  
; Sequence 3, Application US/10699936  
; Publication No. US20050095582A1  
; GENERAL INFORMATION:  
; APPLICANT: Gillim-Ross, Laura  
; APPLICANT: Taylor, Jill  
; APPLICANT: Scholl, David R.  
; APPLICANT: Wentworth, David E.  
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory  
; TITLE OF INVENTION: Syndrome Coronavirus  
; FILE REFERENCE: DHI-07986  
; CURRENT APPLICATION NUMBER: US/10/699,936  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 29736  
; TYPE: DNA  
; ORGANISM: SARS coronavirus CUHK-W1  
US-10-699-936-3

Query Match 100.0%; Score 16; DB 19; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGCGTAGGCAATCC 16  
Db 17745 TCTGCGTAGGCAATCC 17730

Search completed: May 16, 2005, 14:33:47

END PAGE BLANK (uspro)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:01:20 ; Search time 1807 Seconds  
(without alignments)  
337.038 Million cell updates/sec

Title: US-10-808-187A-2473

Perfect score: 16

Sequence: 1 tctgcgtaggcaatcc 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16	100.0	1127	8	CC212165
C 2	16	100.0	1129	5	BU309142
C 3	15	93.8	145	7	CF228802
C 4	15	93.8	379	7	CF588852
C 5	15	93.8	476	6	CB133325
C 6	15	93.8	557	5	BH19258
C 7	15	93.8	560	4	BM302197
C 8	15	93.8	609	7	CK096013
C 9	15	93.8	650	6	CA833034
C 10	15	93.8	667	9	CC821220
C 11	15	93.8	689	6	CA836742
C 12	15	93.8	697	6	CA836130
C 13	15	93.8	726	9	AG437780
C 14	15	93.8	774	2	BE034440
C 15	15	93.8	787	9	AG583208
C 16	14.4	90.0	246	1	AJ692266
C 17	14.4	90.0	284	8	BH017962
C 18	14.4	90.0	347	5	BY321586
C 19	14.4	90.0	384	8	AZ619955
C 20	14.4	90.0	392	7	CF085556
C 21	14.4	90.0	415	1	AV810831
C 22	14.4	90.0	446	2	BF748170
C 23	14.4	90.0	453	5	EX474868
C 24	14.4	90.0	486	2	AW974032

C 25	14.4	90.0	507	2	BF323646
C 26	14.4	90.0	509	4	BI898596
C 27	14.4	90.0	511	2	BF510336
C 28	14.4	90.0	523	1	AA580065
C 29	14.4	90.0	548	8	AZ836547
C 30	14.4	90.0	551	7	CN198330
C 31	14.4	90.0	556	5	BQ039588
C 32	14.4	90.0	563	4	BM176303
C 33	14.4	90.0	564	7	CF248067
C 34	14.4	90.0	576	7	CF247007
C 35	14.4	90.0	580	8	BH518447
C 36	14.4	90.0	582	8	AZ619859
C 37	14.4	90.0	583	5	BP276538
C 38	14.4	90.0	583	5	BP313860
C 39	14.4	90.0	583	5	BP356568
C 40	14.4	90.0	590	6	CB382455
C 41	14.4	90.0	594	4	BJ314417
C 42	14.4	90.0	600	9	CG263972
C 43	14.4	90.0	606	8	AZ579635
C 44	14.4	90.0	621	9	CL345667
C 45	14.4	90.0	629	4	BG687295

## ALIGNMENTS

RESULT 1  
LOCUS CC212165/c 1127 bp DNA linear GSS 12-MAY-2003  
DEFINITION CH261-35M21\_RM1.1 CH261 Gallus gallus genomic clone CH261-35M21,  
genomic survey sequence.

ACCESSION CC212165  
VERSION CC212165  
KEYWORDS GSS.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
REFERENCE 1 (bases 1 to 1127)  
AUTHORS Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE Gallus gallus BAC End Reads  
JOURNAL Unpublished (2003)  
COMMENT Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu

Insert Length: 182000 Std Error: 0.00  
Seq primer: RM1 TAGCACTCACTATAGGAGA  
Class: BAC ends  
High quality sequence start: 113  
High quality sequence stop: 766.

FEATURES  
Location/Qualifiers  
1..1127  
/organism="Gallus gallus"  
/mol\_type="genomic DNA"  
/strain="Red Jungle Fowl"  
/db\_xref="taxon:9031"  
/clone="CH261-35M21"  
/sex="female"  
/cell\_line="UCD001, inbred 256"  
/notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
CH261 Female Chicken library - for library and clone  
ordering information: http://www.chori.org/bacpac"

ORIGIN  
Query Match 100.0%; Score 16; DB 8; Length 1127;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATCC 16



```

asterids; lamids; Gentianales; Rubiaceae; Ixoroideae; Coffeaeae;
Coffea.
REFERENCE
1 (bases 1 to 379)
AUTHORS
Fernandez,D., Santos,P., Agostini,C., Bon,M.C., Petitot,A.S.,
Silva,M.C., Guerra-Guimaraes,L., Ribeiro,A., Argout,X. and
Nicole,M.
TITLE
Coffee (Coffea arabica L.) genes early expressed during infection
by the rust fungus (Hemileia vastatrix)
JOURNAL
Mol. Plant Pathol. 5 (6) (2004) In press
COMMENT
Contact: Fernandez, D.
Equipe Resistance des Plantes aux Parasites, UMR DGPC 1097
IRD
911, avenue Agropolis, BP64501,34394 Montpellier cedex 5, France
Tel: +33 467 416 287
Fax: +33 467 416 283
Email: Diana.Fernandez@mpl.ird.fr
Differentially screened sequence (stronger hybridization signal
with incompatible probe versus compatible probe).
POLYA-No. Location/Qualifiers
FEATURES
source
1. .379
/organism="Coffea arabica"
/mol_type="mRNA"
/cultivar="S4Agaro"
/db_xref="taxon:13443"
/clone="12H9/52"
/lab_host="E.Coli"
/clone_lib="Coffee SSH library 1"
/notes="Organ: leaf; Vector: pGEM-T easy; Clontech
PCR-select cDNA subtraction library; Hypersensitive
reaction (HR)-specific (incompatible interaction as tester
and compatible interaction as driver in the subtraction
process); leaves were collected 12 h post-inoculation with
Hemileia vastatrix"
ORIGIN
Query Match 93.8%; Score 15; DB 7; Length 379;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16
|||||
Db 153 CTGCGTAGGCAATCC 167

RESULT 5
CB133325 476 bp mRNA linear EST 29-JAN-2003
LOCUS
K-EST0184175 L4SNU368s1 Homo sapiens cDNA clone L4SNU368s1-20-H07
DEFINITION
5', mRNA sequence.
ACCESSION
CB133325
VERSION
CB133325.1 GI:28098949
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 476)
AUTHORS
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE
21C Frontier Korean EST Project 2001
JOURNAL
Unpublished (2002)
COMMENT
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yusong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 20 row: H column: 07
High quality sequence stop: 476.
Location/Qualifiers
FEATURES

```

```

1. .476
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L4SNU368s1-20-H07"
/sex="M"
/tissue_type="Liver"
/cell_type="Polygonal"
/cell_line="SNU-368"
/lab_host="Top10F"
/clone_lib="L4SNU368s1"
/notes="Organ: Liver; Vector: pcNS-D2; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 - 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the subtracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F' with
electroporation method."
ORIGIN
Query Match 93.8%; Score 15; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16
|||||
Db 49 CTGCGTAGGCAATCC 63

RESULT 6
BU819258 557 bp mRNA linear EST 15-OCT-2002
LOCUS
UA41BPFO3 Populus tremula cambium cDNA library Populus tremula cDNA
DEFINITION
5 prime, mRNA sequence.
ACCESSION
BU819258
VERSION
BU819258.1 GI:23982788
KEYWORDS
EST.
SOURCE
Populus tremula
ORGANISM
Populus tremula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE
1 (bases 1 to 557)
AUTHORS
Unneberg,P., Bhalerao,R.R., Jansson,S. and Sterky,F.
TITLE
The poplar tree transcriptome: Analysis of expressed sequence tags
from multiple libraries
JOURNAL
Unpublished (2002)
COMMENT
Contact: BHALERAO RUPALI R.
Umea Plant Science Center
Department of Plant Physiology
University of Umea, 901 87 Umea, Sweden

```

Tel: +46 90 786 5279  
 Fax: +46 90 786 6676  
 Email: rupali.bhalerao@plantphys.umu.se.

## FEATURES

source  
 1. .557  
 /organism="Populus tremula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:113636"  
 /tissue\_type="cambium"  
 /clone\_lib="Populus tremula cambium cDNA library"

## ORIGIN

Query Match 93.8%; Score 15; DB 5; Length 557;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGGCTAGGCAATC 15  
 DB 507 TCTGGCTAGGCAATC 521

## RESULT 7

BM302197  
 LOCUS  
 DEFINITION  
 MCA046F07 23578 Ice plant lambda Uni-Zap XR expression library, 0  
 hours NaCl treatment prescreened for removal of highly abundant  
 transcripts Mesembryanthemum crystallinum cDNA clone MCA046F07 5,  
 mRNA sequence.

## ACCESSION

BM302197

## VERSION

BM302197.1 GI:18024572

## KEYWORDS

EST.

## SOURCE

Mesembryanthemum crystallinum (common iceplant)

## REFERENCE

1 (bases 1 to 560)  
 Cushman, J.C.  
 An expressed sequence tag database for the common ice plant,  
 Mesembryanthemum crystallinum  
 Unpublished (1997)

## JOURNAL

Contact: Cushman JC  
 Department of Biochemistry  
 University of Nevada  
 MS200, Reno, NV 89557-0014, USA  
 Tel: 775-784-1918  
 Fax: 775-784-1650

Email: jcushman@unr.edu

PCR Primers

FORWARD: T3 20mer

BACKWARD: T7 21mer

Plate: 046 row: F column: 07

Seq primer: T3 20mer

High quality sequence stop: 560.

## FEATURES

source  
 1. .560  
 /organism="Mesembryanthemum crystallinum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3544"  
 /clone="MCA046F07"  
 /tissue\_type="leaf"  
 /dev\_stage="six-week-old"  
 /clone\_lib="Ice plant lambda Uni-Zap XR expression  
 library, 0 hours NaCl treatment prescreened for removal of  
 highly abundant transcripts"  
 /note="Vector: lambda Uni-Zap XR, Bluescript SK-; Site\_1:  
 EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 93.8%; Score 15; DB 4; Length 560;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16  
 DB 12 CTGCGTAGGCAATCC 26

## RESULT 8

CK096013/c

LOCUS

DEFINITION

CK096013

ACCESSION

CK096013

VERSION

CK096013.1 GI:38580338

KEYWORDS

SOURCE

ORGANISM

Populus tremula

Populus tremula

REFERENCE

AUTHORS

1 (bases 1 to 609)

Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P.,

Brunner, A.M., Campaa, L., Jonsson-Lindvall, J., Tandré, K.,

Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P.,

Nilsson, O., Sandberg, G., Karlsson, J., Lundberg, J. and Jansson, S.

A Populus EST resource for functional genomics

Unpublished (2003)

Other ESTs: UA41BPF03, UA41BPF03.5pr

Contact: Bo Segerman

Umea Plant Science Center, Department of Plant Physiology

Umea University

901 87 Umea, Sweden

Tel: +46 90 786 5279

Fax: +46 90 786 6676

Email: bo.segerman@plantphys.umu.se.

Location/Qualifiers

1. .609

/organism="Populus tremula"

/mol\_type="mRNA"

/db\_xref="taxon:113636"

/clone="UA41BPF03"

/tissue\_type="Dormant cambium"

/clone\_lib="Populus dormant cambium cDNA library"

ORIGIN

Query Match 93.8%; Score 15; DB 7; Length 609;

Best Local Similarity 100.0%; Pred. No. 7.5e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGGCTAGGCAATC 15  
 DB 102 TCTGGCTAGGCAATC 88

## RESULT 9

CA833034

LOCUS

DEFINITION

CA833034

ACCESSION

CA833034

VERSION

CA833034.1 GI:26560799

KEYWORDS

EST.

SOURCE

Mesembryanthemum crystallinum (common iceplant)

Mesembryanthemum crystallinum

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Aizoaceae; Mesembryanthemum.

1 (bases 1 to 650)

Cushman, J.C.

An expressed sequence tag database for the common ice plant,

Mesembryanthemum crystallinum

Unpublished (1997)

JOURNAL

Contact: Cushman JC

CK096013 609 bp mRNA linear EST 01-DEC-2003  
 UA41BPF03.3pr Populus dormant cambium cDNA library Populus tremula  
 cDNA clone UA41BPF03 3', mRNA sequence.

CK096013

CK096013.1 GI:38580338

EST.

Populus tremula

Populus tremula

REFERENCE

1 (bases 1 to 609)

Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P.,

Brunner, A.M., Campaa, L., Jonsson-Lindvall, J., Tandré, K.,

Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P.,

Nilsson, O., Sandberg, G., Karlsson, J., Lundberg, J. and Jansson, S.

A Populus EST resource for functional genomics

Unpublished (2003)

Other ESTs: UA41BPF03, UA41BPF03.5pr

Contact: Bo Segerman

Umea Plant Science Center, Department of Plant Physiology

Umea University

901 87 Umea, Sweden

Tel: +46 90 786 5279

Fax: +46 90 786 6676

Email: bo.segerman@plantphys.umu.se.

Location/Qualifiers

1. .609

/organism="Populus tremula"

/mol\_type="mRNA"

/db\_xref="taxon:113636"

/clone="UA41BPF03"

/tissue\_type="Dormant cambium"

/clone\_lib="Populus dormant cambium cDNA library"

CA833034 650 bp mRNA linear EST 12-DEC-2002  
 MCS019C05.151346 Ice plant lambda Uni-Zap XR expression library, 5  
 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase 1 (2  
 AM). Mesembryanthemum crystallinum cDNA clone MCS019C05 5, mRNA  
 sequence.

CA833034

CA833034.1 GI:26560799

EST.

Mesembryanthemum crystallinum (common iceplant)

Mesembryanthemum crystallinum

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Aizoaceae; Mesembryanthemum.

1 (bases 1 to 650)

Cushman, J.C.

An expressed sequence tag database for the common ice plant,

Mesembryanthemum crystallinum

Unpublished (1997)

JOURNAL

Contact: Cushman JC

Department of Biochemistry  
University of Nevada  
MS200, Reno, NV 89557-0014, USA  
Tel: 775-784-1918  
Fax: 775-784-1650  
Email: jcushman@unr.edu  
PCR Primers  
FORWARD: T3 20mer  
BACKWARD: T7 21mer  
Plate: 019 row: C column: 05  
Seq primer: T3 20mer  
High quality sequence stop: 650.  
Location/Qualifiers

## FEATURES

source

1. .650  
/organism="Mesembryanthemum crystallinum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3544"  
/clone="MCS019C05"  
/tissue\_type="leaf"  
/dev\_stage="five-week-old"  
/clone\_lib="Ice plant Lambda Uni-Zap XR expression library, 5 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase I (2 AM)."  
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Library construction was performed according to Stratagene's recommended protocol for the Lambda UniZapXR vector and cDNA synthesis kit."

## ORIGIN

Query Match 93.8%; Score 15; DB 6; Length 650;  
Best Local Similarity 100.0%; Pred. No. 7.5e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16  
|||||  
Db 385 CTGCGTAGGCAATCC 399

## RESULT 10

CC821220

LOCUS

DEFINITION 667 bp DNA linear GSS 22-OCT-2003  
PECP6A8 Uncultured human fecal virus uncultured human fecal virus  
genomic, genomic survey sequence.

ACCESSION

CC821220

VERSION

GI:37807020

KEYWORDS

uncultured human fecal virus

SOURCE

Viruses; environmental samples.

ORGANISM

1 (bases 1 to 667)

REFERENCE

Breitbart, M., Hewson, I., Felts, B., Mahaffy, J.M., Nulton, J.,

Salamon, P. and Rohwer, F.

Metagenomic analyses of an uncultured viral community from human

feces

J. Bacteriol. 185 (20), 6220-6223 (2003)

J. Bacteriol. 185 (20), 6220-6223 (2003)

MEDLINE

22887472

PUBMED

14526037

COMMENT

Contact: Rohwer F

Biology Dept.

San Diego State University

5500 Campanile Dr, San Diego, CA 92102, USA

Tel: 6195941336

Fax: 619595676

Email: forest@unstroke.sdsu.edu

Class: shotgun.

Location/Qualifiers

1. .667

## FEATURES

source

/organism="uncultured human fecal virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:239364"  
/clone\_lib="Uncultured human fecal virus"

## ORIGIN

Query Match 93.8%; Score 15; DB 6; Length 689;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16

|||||

Db 12 CTGCGTAGGCAATCC 26

RESULT 12

CA836130

LOCUS

DEFINITION

CA836130

697 bp mRNA linear EST 12-DEC-2002

MCU003H04\_254822 Ice plant Lambda Uni-Zap XR expression library, 5

days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase II

Query Match 93.8%; Score 15; DB 9; Length 667;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATC 15  
|||||

Db 465 TCTGCGTAGGCAATC 479

## RESULT 11

CA836742

LOCUS

DEFINITION

CA836742

689 bp mRNA linear EST 12-DEC-2002

MCU010E08\_256046 Ice plant Lambda Uni-Zap XR expression library, 5

days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase II

(8:00 AM). Mesembryanthemum crystallinum cDNA clone MCU010E08 5,

mRNA sequence.

ACCESSION

CA836742

VERSION

GI:26564507

KEYWORDS

EST.

SOURCE

Mesembryanthemum crystallinum (common iceplant)

ORGANISM

Mesembryanthemum crystallinum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Aizoaceae; Mesembryanthemum.

REFERENCE

1 (bases 1 to 689)

Cushman, J.C.

An expressed sequence tag database for the common ice plant,

Mesembryanthemum crystallinum

Unpublished (1997)

COMMENT

Contact: Cushman JC

Department of Biochemistry

University of Nevada

MS200, Reno, NV 89557-0014, USA

Tel: 775-784-1918

Fax: 775-784-1650

Email: jcushman@unr.edu

PCR Primers

FORWARD: T3 20mer

BACKWARD: T7 21mer

Plate: 010 row: E column: 08

Seq primer: T3 20mer

High quality sequence stop: 689.

Location/Qualifiers

1. .689

/organism="Mesembryanthemum crystallinum"

/mol\_type="mRNA"

/db\_xref="taxon:3544"

/clone="MCU010E08"

/tissue\_type="leaf"

/dev\_stage="five-week-old"

/clone\_lib="Ice plant Lambda Uni-Zap XR expression

library, 5 days 0.5 M NaCl treatment, Crassulacean acid

metabolism, phase II (8:00 AM)."

/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:

EcoRI; Site 2: XhoI; Library construction was performed

according to Stratagene's recommended protocol for the

Lambda UniZapXR vector and cDNA synthesis kit."

ORIGIN

Query Match

Best Local Similarity

Matches

15; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

2

CTGCGTAGGCAATCC

16

|||||

Db

12

CTGCGTAGGCAATCC

26

RESULT 12

CA836130

LOCUS

DEFINITION

CA836130

697 bp mRNA linear EST 12-DEC-2002

MCU003H04\_254822 Ice plant Lambda Uni-Zap XR expression library, 5

days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase II

```

(8:00 AM). Mesembryanthemum crystallinum cDNA clone MCU003H04 5,
mRNA sequence.
ACCESSION   CA836130.1   GI:26563895
VERSION     CA836130.1   GI:26563895
KEYWORDS    EST.
SOURCE      Mesembryanthemum crystallinum (common iceplant)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllales; Alzooaceae; Mesembryanthemum.
REFERENCE   1 (bases 1 to 697)
AUTHORS    Cushman, J.C.
TITLE      An expressed sequence tag database for the common ice plant,
            Mesembryanthemum crystallinum
JOURNAL     Unpublished (1997)
COMMENT     Contact: Cushman JC
            Department of Biochemistry
            University of Nevada
            MS200, Reno, NV 89557-0014, USA
            Tel: 775-784-1918
            Fax: 775-784-1650
            Email: jcushman@unr.edu
PCR PRIMERS
            FORWARD: T3 20mer
            BACKWARD: T7 21mer
Plate: 003 row: H column: 04
Seq primer: T3 20mer
High quality sequence stop: 697.
            Location/Qualifiers
FEATURES   source
            1..697
            /organism="Mesembryanthemum crystallinum"
            /mol_type="mRNA"
            /db_xref="taxon:3544"
            /clone="MCU003H04"
            /tissue_type="leaf"
            /dev_stage="five-week-old"
            /clone_lib="Ice plant Lambda Uni-Zap XR expression
            library, 5 days 0.5 M NaCl treatment, Crassulacean acid
            metabolism, phase II (8:00 AM)."
            /note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:
            EcoRI; Site 2: XhoI; Library construction was performed
            according to Stratagene's recommended protocol for the
            Lambda UniZapXR vector and cDNA synthesis kit."
ORIGIN
Query Match      93.8%; Score 15; DB 6; Length 697;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 CTGCGTAGGCAATCC 16
        |||||
        12 CTGCGTAGGCAATCC 26

RESULT 13
AG437780
LOCUS      AG437780          726 bp DNA linear GSS 03-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-317M24.TJ, genomic survey
            sequence.
ACCESSION   AG437780
VERSION     AG437780.1   GI:48080843
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS    Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE      BAC end Sequences of Library MSMg01
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 726)
AUTHORS    Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE      Direct Submission

```

```

Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@psc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@tc.riken.jp).
Tsukuba Institute, Bio Resource Center.
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@tc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY      Vector : pBACe3.6
            R.Site 1 : EcoRI
            R.Site 2 : EcoRI.
FEATURES     Location/Qualifiers
            1..726
            /organism="Mus musculus molossinus"
            /mol_type="genomic DNA"
            /sub_species="molossinus"
            /db_xref="taxon:57486"
            /clone="MSMg01-317M24.TJ"
            /sex="male"
            /tissue_type="mixture of kidney and spleen"
            /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match      93.8%; Score 15; DB 9; Length 726;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCTGCGTAGGCAATC 15
        |||||
        655 TCTGCGTAGGCAATC 669

RESULT 14
BE034440
LOCUS      BE034440          774 bp mRNA linear EST 07-JUN-2000
DEFINITION MH04H06 MH Mesembryanthemum crystallinum cDNA 5' similar to
            translation initiation factor eif-4a.9, mRNA sequence.
ACCESSION   BE034440
VERSION     BE034440.1   GI:8329449
KEYWORDS    EST.
SOURCE      Mesembryanthemum crystallinum (common iceplant)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllales; Alzooaceae; Mesembryanthemum.
REFERENCE   1 (bases 1 to 774)
AUTHORS    Bohnert, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M.,
            Ferrea, H., Kawasaki, S., McCollough, A., Michalowski, C.B.,
            Palacio, C., Scara, G., Wheeler, M. and Zepeda, G.R.
            Functional Genomics of Plant Stress Tolerance
            Unpublished (2000)
            Contact: Michalowski, C.B.
            University of Arizona
            Bio Sciences West room 513, Tucson, AZ 85721, USA
            Tel: 520-621-7982
            Fax: 520-621-1697
            Email: cbm@u.arizona.edu
            An open reading frame exists.
            Insert Length: 1 Std Error: 0.00.
FEATURES     Location/Qualifiers
            1..774
            /organism="Mesembryanthemum crystallinum"
            /mol_type="mRNA"
            /db_xref="taxon:3544"
            /tissue_type="roots"
            /dev_stage="5-6 weeks"

```

Job time : 1815 secs

/clone\_lib="MH"  
/note="grown in hydroponics, stress 400 mM NaCl (in 0.5  
Hoagland's), 78 h stress"

ORIGIN

Query Match 93.8%; Score 15; DB 2; Length 774;  
Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTGCGTAGGCAATCC 16

|||||

Db 419 CTGCGTAGGCAATCC 433

RESULT 15

AG583208

LOCUS

DEFINITION

AG583208 787 bp DNA linear GSS 05-JUN-2004  
Mus musculus molossinus DNA, clone:MSMg01-509123.TJ, genomic survey  
sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AG583208 GI:48344038

GSS.

Mus musculus molossinus

Mus musculus molossinus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.

BAC end Sequences of Library MSMg01

1

2 (bases 1 to 787)

Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.

Direct Submission

Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the mouse BAC library MSMg01. For BAC

library availability, please contact Kuniya Abe (abe@rtc.riken.jp).

Tsukuba Institute, Bio Resource Center.

The Institute of Physical and Chemical Research (RIKEN) 3-1-1

Koyadai, Tsukuba, 305-0074 Japan

phone: 81-298-36-9189, fax: 81-298-36-9199

e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : TJ

LIBRARY

Vector : pBACe3.6

R.Site 1 : ECGRI

R.Site 2 : ECGRI

FEATURES

source

1..787 Location/Qualifiers

/organism="Mus musculus molossinus"

/mol\_type="genomic DNA"

/sub\_species="molossinus"

/db\_xref="taxon:57486"

/clone="MSMg01-509123.TJ"

/sex="male"

/tissue\_type="mixture of kidney and spleen"

/clone\_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 93.8%; Score 15; DB 9; Length 787;  
Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGCGTAGGCAATC 15

|||||

Db 654 TCTGCGTAGGCAATC 668

Search completed: May 16, 2005, 06:03:28

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 21:16:45 ; Search time 502.031 Seconds  
(without alignments)  
2026.885 Million cell updates/sec

Title: US-10-808-187A-2474

Perfect score: 21  
Sequence: 1 accagaatgagagcaatg 21

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	1269	14	AY307165 SARS coro
2	21	100.0	1269	14	AY360146 SARS coro
3	21	100.0	1269	14	AY365036 SARS coro
4	21	100.0	1669	14	AY536760 SARS coro
5	21	100.0	1873	14	AY534758S4
6	21	100.0	1938	14	AY534762S6
7	21	100.0	2304	14	AY322205S4
8	21	100.0	2810	14	AY290752 SARS coro
9	21	100.0	3080	14	AY443086S10
10	21	100.0	8439	14	AY304489 SARS coro
11	21	100.0	8581	14	AY304487 SARS coro
12	21	100.0	11006	14	AY304491 SARS coro
13	21	100.0	11010	14	AY304493 SARS coro
14	21	100.0	11010	14	AY304494 SARS coro
15	21	100.0	13471	14	AY304490 SARS coro
16	21	100.0	13471	14	AY304492 SARS coro
17	21	100.0	29013	14	AY463060 SARS coro
18	21	100.0	29350	14	AY394999 SARS coro
19	21	100.0	29350	14	AY395000 SARS coro

20	21	100.0	29350	14	AY395001 SARS coro
21	21	100.0	29350	14	AY395002 SARS coro
22	21	100.0	29433	14	AY394977 SARS coro
23	21	100.0	29530	14	AY394985 SARS coro
24	21	100.0	29573	14	AY338174 SARS coro
25	21	100.0	29573	14	AY338175 SARS coro
26	21	100.0	29573	14	AY348314 SARS coro
27	21	100.0	29577	14	AY559094 SARS coro
28	21	100.0	29592	14	AY463059 SARS coro
29	21	100.0	29620	14	AY395004 SARS coro
30	21	100.0	29640	14	AY394978 SARS coro
31	21	100.0	29645	14	AY394979 SARS coro
32	21	100.0	29646	14	AY394982 SARS coro
33	21	100.0	29647	14	AY395003 SARS coro
34	21	100.0	29661	14	AY559086 SARS coro
35	21	100.0	29665	14	AY394988 SARS coro
36	21	100.0	29670	14	AY559082 SARS coro
37	21	100.0	29683	14	AY394996 SARS coro
38	21	100.0	29683	14	AY394997 SARS coro
39	21	100.0	29699	14	AY394983 SARS coro
40	21	100.0	29705	14	AY283795 SARS coro
41	21	100.0	29705	14	AY394980 SARS coro
42	21	100.0	29706	14	AY283797 SARS coro
43	21	100.0	29709	14	AY394987 SARS coro
44	21	100.0	29710	14	AY559091 SARS coro
45	21	100.0	29711	14	AY283794 SARS coro

ALIGNMENTS

RESULT 1  
AY307165  
LOCUS 1269 bp RNA linear VRL 09-JUN-2003  
DEFINITION SARS coronavirus nucleocapsid protein (NP) gene, complete cds.  
ACCESSION AY307165  
VERSION AY307165.1 GI:31540948  
KEYWORDS  
SOURCE SARS coronavirus  
ORGANISM SARS coronavirus  
VIRUSES: ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus.  
REFERENCE 1 (bases 1 to 1269)  
AUTHORS Sun, K., Anwar, A., Gupta, V., Tablin, M.T., Atkinson, R., Chandrasekarn, A., and August, T.J.  
TITLE Direct Submission  
JOURNAL Submitted (26-MAY-2003) Genetic Immunotherapy Lab, Johns Hopkins Singapore, 41 Science Park Road, #03-18 The Gemini, Singapore, Singapore 117610, Singapore  
FEATURES  
source Location/Qualifiers  
1..1269  
/organism="SARS coronavirus"  
/virion  
/mol\_type="genomic RNA"  
/db\_xref="taxon:227859"  
gene 1..1269  
/gene="NP"  
CDS 1..1269  
/gene="NP"  
/note="associated with virus genomic RNA; involved in CTL response"  
/codon\_start=1  
/product="nucleocapsid protein"  
/protein\_id="AAP49024.1"  
/db\_xref="GI:31540949"  
/translation="MSDNGPQSNQRSAPIITFGGPTDSTNNQNGNGARPKQRFPQ  
GLPNNTASWFTALTQHKELRFPKQGVINTNSGDDQIGIYTRVNRGGDKM  
KELSPWYFYLTGTGPEASLPYGANKGIVWVATEGALNPKDHIGTRNPNNAATVL  
ALPQGTTLPRGFVAEGSGGSOASSSSRSRNSRSTPGSSRGNSPARMASGGGT  
OLALLLLDNLNOLSKSVGGQOQGVTTKKSAARSKKPKROKRTATKYNTVQAFG  
RRGPEOTQGNFGQDQLIRQGTDTYKHPQIAQFAPSASAFEGMSRIGNEVTPSGTWLY  
HGAIKLDDKDPQFKDNLINKHIDAYKTFFPTPEPKKKKKTKDEAQLPQKQKQPT  
VTLLPAADMDDFSRQLQNSMSGASADSTQA"

```

ORIGIN
Query Match      100.0%; Score 21; DB 14; Length 1269;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGACGCAATG 21
    |||||||||||||||||||
Db 83 ACCAGAATGGAGGACGCAATG 103

RESULT 2
AY360146
LOCUS SARS coronavirus HPZ-2003 linear VRL 02-SEP-2003
DEFINITION SARS coronavirus HPZ-2003 nucleocapsid protein (NP) gene, complete cds.
ACCESSION AY360146
VERSION AY360146.1 GI:34329618
KEYWORDS SARS coronavirus HPZ-2003
SOURCE SARS coronavirus HPZ-2003
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 1269)
AUTHORS Zhu,H.P., Lu,Q.Y., Lu,Y.Y., Yao,P.P., Xu,F., Ge,Q., Weng,J.Q., Yan,J.Y., Gong,L.M., Shi,W., Zhao,Z.Y. and Zhu,Z.Y.
TITLE Cloning and expression of nucleocapsid protein gene of SARS associated coronavirus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1269)
AUTHORS Zhu,H.P., Lu,Q.Y., Lu,Y.Y., Yao,P.P., Xu,F., Ge,Q., Weng,J.Q., Yan,J.Y., Gong,L.M., Shi,W., Zhao,Z.Y. and Zhu,Z.Y.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-2003) Zhejiang Provincial Key Laboratory of HFRS, Zhejiang Center for Disease Prevention and Control, 17 Lao Zhe Da Zhi Road, Hangzhou, Zhejiang 310009, P.R. China
FEATURES
    source
        1..1269
            /organism="SARS coronavirus HPZ-2003"
            /mol_type="genomic RNA"
            /db_xref="taxon:243925"
            1..1269
            /gene="NP"
            1..1269
            /gene="NP"
            /codon_start=1
            /product="nucleocapsid protein"
            /protein_id="AAQ63890.1"
            /db_xref="GI:34329619"
            /translation="MSDNGPQSNRSAPRITFGGPTDSTDNNGRNGARPQKRRPQ
            GLPNTASWFTALTQHGKEELRFRGQGVPIINTSGPDDQIGYYRATRRVGGGDKM
            KELSFRWYFYLLGTGPEASLPYGANKGIVWVATEGALNTPKDHIGTRNPNNAATVL
            QLPQGTTLPGFYAEGSRGGSQASRSRSGNSRSTPGSSRSGNSPARMSGGGET
            ALALLDLRLNLESKVGSKGQQGQQTIVTKKSAEASKKPRKRTATKQYNVTQAFG
            RRGPEQTQGNFGDQLIRQGTDKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY
            HGAIKLDKDPQKDNVILNKHIDAYKTFPTPEPKDKKKTKDEAQLPQKQKQPT
            VTLLPAADMDDFSRLQNSMSGASADSTQA"

ORIGIN
Query Match      100.0%; Score 21; DB 14; Length 1269;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGACGCAATG 21
    |||||||||||||||||||
Db 83 ACCAGAATGGAGGACGCAATG 103

RESULT 3
AY365036
LOCUS SARS coronavirus HB linear VRL 01-JUN-2004
DEFINITION SARS coronavirus HB nucleocapsid protein (NP) gene, complete cds.
ACCESSION AY365036

```

```

VERSION AY365036.1 GI:38176101
KEYWORDS SARS coronavirus HB
SOURCE SARS coronavirus HB
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 1269)
AUTHORS Timani,K.A., Ye,L., Ye,L., Zhu,Y., Wu,Z. and Gong,Z.
TITLE Cloning, sequencing, expression, and purification of SARS-associated coronavirus nucleocapsid protein for serodiagnosis of SARS
JOURNAL J. Clin. Virol. 30 (4), 309-312 (2004)
PUBMED 15163419
REFERENCE 2 (bases 1 to 1269)
AUTHORS Timani,K.A., Ye,L., Ye,L., Zhu,Y. and Guo,D.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-2003) Institute of Virology, Wuhan University, Wuhan, Hubei 430072, P.R. China
FEATURES
    Location/Qualifiers
        1..1269
            /organism="SARS coronavirus HB"
            /viral
            /mol_type="genomic RNA"
            /isolate="HB"
            /db_xref="taxon:253634"
            /country="China: Wuhan, Hubei"
            1..1269
            /gene="NP"
            1..1269
            /gene="NP"
            /note="N protein"
            /codon_start=1
            /product="nucleocapsid protein"
            /protein_id="AARI2990.1"
            /db_xref="GI:38176102"
            /translation="MSDNGPQSNRSAPRITFGGPTDSTDNNGRNGARPQKRRPQ
            GLPNTASWFTALTQHGKEELRFRGQGVPIINTSGPDDQIGYYRATRRVGGGDKM
            KELSFRWYFYLLGTGPEASLPYGANKGIVWVATEGALNTPKDHIGTRNPNNAATVL
            QLPQGTTLPGFYAEGSRGGSQASRSRSGNSRSTPGSSRSGNSPARMSGGGET
            ALALLDLRLNLESKVGSKGQQGQQTIVTKKSAEASKKPRKRTATKQYNVTQAFG
            RRGPEQTQGNFGDQLIRQGTDKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY
            HGAIKLDKDPQKDNVILNKHIDAYKTFPTPEPKDKKKTKDEAQLPQKQKQPT
            VTLLPAADMDDFSRLQNSMSGASADSTQA"

ORIGIN
Query Match      100.0%; Score 21; DB 14; Length 1269;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGACGCAATG 21
    |||||||||||||||||||
Db 83 ACCAGAATGGAGGACGCAATG 103

RESULT 4
AY536760
LOCUS SARS coronavirus BJ01 linear VRL 15-SEP-2004
DEFINITION SARS coronavirus BJ01 nucleocapsid protein mRNA, complete cds.
ACCESSION AY536760
VERSION AY536760.3 GI:52100973
KEYWORDS SARS coronavirus BJ01
SOURCE SARS coronavirus BJ01
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 1669)
AUTHORS Li,T., Li,X., Chang,Z. and Liu,L.
TITLE Identification of SARS-CoV mRNA leader sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1669)
AUTHORS Li,T., Li,X., Liu,L. and Chang,Z.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2004) Institute of Biomedicine,Tsinghua

```

REFERENCE  
AUTHORS University, Beijing, Beijing 100084, China  
TITLE 3 (bases 1 to 1669)  
JOURNAL Li, T., Li, X., Liu, L. and Chang, Z.  
REMARK Direct Submission  
COMMENT Submitted (06-JUL-2004) Institute of Biomedicine, Tsinghua  
AUTHORS University, Beijing, Beijing 100084, China  
TITLE Sequence update by submitter  
JOURNAL 4 (bases 1 to 1669)  
REMARK Li, T., Li, X., Chang, Z. and Liu, L.  
AUTHORS Direct Submission  
TITLE Submitted (15-SEP-2004) Institute of Biomedicine, Tsinghua  
JOURNAL University, Beijing 100084, China  
REMARK Sequence update by submitter  
COMMENT On Sep 15, 2004 this sequence version replaced gi:49921010.  
FEATURES  
source Location/Qualifiers  
1. .1669  
/organism="SARS coronavirus BJ01"  
/mol\_type="mRNA"  
/strain="BJ01"  
/db\_xref="taxon:228407"  
/country="China"  
81. .1349  
/codon\_start=1  
/product="nucleocapsid protein"  
/protein\_id="AA848456.1"  
/db\_xref="GI:44889746"  
/translation="MSDNGPQNSRAPSRIITFGGPTDSTNNQNGRNGARPKORRPP  
GLPNFTASFTALTQKEELRPRGGVPEINTNSGPDQIGYIRTRVRGGDGKM  
KELSPWYFYLGTGPEALPYGANKKEGIVWVATEGALNTPKDHI GTRPNNAATVL  
ALPQGTTLPRGFVAEGSRGGSQASRSSRSRGNSTPGSSRSGNSPARMSGGET  
OLALLLDLNLQLESKVSQGGQOQGVTTKKSAAEASKPKRKTATKYNYTQAFG  
RRPEQTQGNFGQDLIRQGTDTYHFWFQIAQFAPSAPFGMSKRI GHEVTPSTWLT  
VGAIKDDKDPQPNVLLNKHIDAYKTFPTEPKKKKKTKDEAQLPQPKKKQFT  
VTPLPAADWMDDFSRQLQNSMSGASADSTQA"

ORIGIN  
Query Match 100.0%; Score 21; DB 14; Length 1669;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCAGAATGGAGCGCAATG 21  
|||||  
Db 163 ACCAGAATGGAGCGCAATG 183  
|||||

RESULT 5  
AY534758S4  
LOCUS 1873 bp RNA linear VRL 17-MAR-2004  
DEFINITION SARS coronavirus Sin0409, partial sequence.  
ACCESSION AY534761  
VERSION AY534761.1 GI:45384968  
KEYWORDS 4 of 4  
SEGMENT SARS coronavirus Sin0409  
SOURCE SARS coronavirus Sin0409  
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.  
REFERENCE 1 (bases 1 to 1873)  
AUTHORS Lim, P.L., Kurup, A., Gopalakrishna, G., Chan, K.P., Wong, C.W.,  
Ng, L.C., Se-Thoe, S.Y., Oon, L., Bai, X., Stanton, L.W., Ruan, Y.,  
Miller, L.D., Vega, V.B., James, L., Ooi, P.L., Kai, C.S., Olsen, S.J.,  
Ang, B. and Leo, Y.S.  
TITLE Laboratory-acquired severe acute respiratory syndrome (SARS) -  
Singapore 2003  
JOURNAL Location/Qualifiers  
1. .1873  
/organism="SARS coronavirus Sin0409"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:266148"  
/country="Singapore"

ORIGIN  
Query Match 100.0%; Score 21; DB 14; Length 1938;  
Best Local Similarity 100.0%; Pred. No. 7.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCAGAATGGAGCGCAATG 21  
|||||  
Db 428 ACCAGAATGGAGCGCAATG 448  
|||||

RESULT 7  
AY322205S4  
LOCUS 2304 bp RNA linear VRL 21-JUL-2003  
DEFINITION SARS coronavirus Shanghai LY Orf7a gene, partial cds; and Orf7b,  
Orf8a, Orf8b, and nucleocapsid protein genes, complete cds.  
ACCESSION AY322208  
VERSION AY322208.1 GI:32454342  
KEYWORDS 4 of 4  
SEGMENT SARS coronavirus Shanghai LY  
SOURCE SARS coronavirus Shanghai LY  
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.  
REFERENCE 1 (bases 1 to 2304)  
AUTHORS Yuan, Z., Zhang, X., Hu, Y., Lan, S., Wang, H., Zhou, Z. and Wen, Y.

ORIGIN  
Query Match 100.0%; Score 21; DB 14; Length 1873;  
Best Local Similarity 100.0%; Pred. No. 7.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCAGAATGGAGCGCAATG 21  
|||||  
Db 402 ACCAGAATGGAGCGCAATG 422  
|||||

RESULT 6  
AY534762S6  
LOCUS 1938 bp RNA linear VRL 17-MAR-2004  
DEFINITION SARS coronavirus Sin\_WNV, partial sequence.  
ACCESSION AY534767  
VERSION AY534767.1 GI:45384975  
KEYWORDS 6 of 6  
SEGMENT SARS coronavirus Sin\_WNV  
SOURCE SARS coronavirus Sin\_WNV  
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.  
REFERENCE 1 (bases 1 to 1938)  
AUTHORS Lim, P.L., Kurup, A., Gopalakrishna, G., Chan, K.P., Wong, C.W.,  
Ng, L.C., Se-Thoe, S.Y., Oon, L., Bai, X., Stanton, L.W., Ruan, Y.,  
Miller, L.D., Vega, V.B., James, L., Ooi, P.L., Kai, C.S., Olsen, S.J.,  
Ang, B. and Leo, Y.S.  
TITLE Laboratory-acquired severe acute respiratory syndrome (SARS) -  
Singapore 2003  
JOURNAL Location/Qualifiers  
1. .1938  
/organism="SARS coronavirus Sin\_WNV"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:266148"  
/country="Singapore"

ORIGIN  
Query Match 100.0%; Score 21; DB 14; Length 1938;  
Best Local Similarity 100.0%; Pred. No. 7.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCAGAATGGAGCGCAATG 21  
|||||  
Db 428 ACCAGAATGGAGCGCAATG 448  
|||||

RESULT 7  
AY322205S4  
LOCUS 2304 bp RNA linear VRL 21-JUL-2003  
DEFINITION SARS coronavirus Shanghai LY Orf7a gene, partial cds; and Orf7b,  
Orf8a, Orf8b, and nucleocapsid protein genes, complete cds.  
ACCESSION AY322208  
VERSION AY322208.1 GI:32454342  
KEYWORDS 4 of 4  
SEGMENT SARS coronavirus Shanghai LY  
SOURCE SARS coronavirus Shanghai LY  
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.  
REFERENCE 1 (bases 1 to 2304)  
AUTHORS Yuan, Z., Zhang, X., Hu, Y., Lan, S., Wang, H., Zhou, Z. and Wen, Y.

TITLE Direct Submission  
JOURNAL Submitted (12-JUN-2003) Molecular Virology, Shanghai Medical College of Fudan University, 138 Yi Xue Yuan Road, Shanghai 200032, P.R. China

FEATURES

source Location/Qualifiers  
1..2304  
/organism="SARS coronavirus Shanghai LY"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:235173"  
/country="China: Shanghai"  
<1..276  
/codon\_start=1  
/product="Orf7a"  
/protein\_id="AAP82970.1"  
/db\_xref="GI:32454348"  
/translation="KEPCPSATYEGNSPFPLADNKKFALTCTSTHFACADGTRHTY  
OLRARSVPKLFIRQEEVQELYSPLFLVLAALVFLILCFTIKRTE"  
273..407

CDS

<1..276  
/codon\_start=1  
/product="Orf7a"  
/protein\_id="AAP82970.1"  
/db\_xref="GI:32454348"  
/translation="KEPCPSATYEGNSPFPLADNKKFALTCTSTHFACADGTRHTY  
OLRARSVPKLFIRQEEVQELYSPLFLVLAALVFLILCFTIKRTE"  
273..407

CDS

/codon\_start=1  
/product="Orf7a"  
/protein\_id="AAP82971.1"  
/db\_xref="GI:32454349"  
/translation="MNEULTIDFYLCFLAFLFLVLIIMLIIFWFSLEIQDLEPCTKV  
"  
414..533

CDS

/codon\_start=1  
/product="Orf8a"  
/protein\_id="AAP82972.1"  
/db\_xref="GI:32454350"  
/translation="MKLLIVLTICISLCICITVQVCASNKPVLDEPCKVQH"  
499..753

CDS

/codon\_start=1  
/product="Orf8b"  
/protein\_id="AAP82973.1"  
/db\_xref="GI:32454351"  
/translation="MCKLILVYNTRGNTYSTAWLCALGKVLPHFRHWTMTVQCTPNV  
TINCQDPAGGALLIARCWLHGHQTAAPRDVLVLNKTN"  
755..2023

CDS

/codon\_start=1  
/product="nucleocapsid protein"  
/protein\_id="AAP82974.1"  
/db\_xref="GI:32454352"  
/translation="MSDNGPQSNQRSAPRIFFGGTSDTNNQNGRNGARPKQRRPQ  
GLPNNATFWTALTQHGKEELRFPFGQGVPIINTNSGDDQIGYRRATRRVRGGDKM  
KELSPRTWYFYLGTGPEASLPGYANKGIVWVATEGALNTKDHIGTRPNPNNATVL  
QLPQGTTLPLKGFYAGSRGSGQASRSRSGNSRSTPGSSRSGNSPARMASGGGT  
ALALLDLNLESKVGKGGQGGQGTWTKSAEASKKPKRKTATKQYNVTQAFG  
RPGPQTGNGFDQDLIRQGTQYKHPQIAOPAPSASAFFGMSRIGMEAPSGTWLTY  
HGAIKDDKDPQKONVLLNKHIDAYKTFFPTEPKDKKKKTDEAQLPQKQKQPT  
VTLLPAADMDDFSRQLQNSMGASADSTQA"

ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 2304;  
Best Local Similarity 100.0%; Pred. No. 7.7; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACCAGAATGAGGACGCAATG 21  
|||||  
Db 837 ACCAGAATGAGGACGCAATG 857  
|||||

RESULT 8

AY290752

LOCUS

DEFINITION SARS coronavirus ZJ01 isolate 2810 bp RNA linear VRL 09-FEB-2004  
uncharacterized protein 7a, uncharacterized protein 7b,  
uncharacterized protein 8a, uncharacterized protein 8b,  
nucleocapsid protein, uncharacterized protein 9b, and  
uncharacterized protein 9c genes, complete cds.

AY290752

ACCESSION

AY290752.2

VERSION

KEYWORDS

SOURCE

ORGANISM

SARS coronavirus ZJ01  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE

1 (bases 1 to 2810)  
Li, L., Wang, Z., Lu, Y., Bao, Q., Chen, S., Wu, N., Cheng, S., Weng, J.,  
Zhang, Y., Yan, J., Mei, L., Wang, X., Zhu, H., Yu, Y., Zhang, M., Li, M.,  
Yao, J., Lu, Q., Yao, P., Bo, X., Wo, J., Wang, S. and Hu, S.  
Severe acute respiratory syndrome-associated coronavirus genotype  
and its characterization  
Chin. Med. J. 116 (9), 1288-1292 (2003)  
22889812

TITLE

Severe acute respiratory syndrome-associated coronavirus genotype  
and its characterization  
Chin. Med. J. 116 (9), 1288-1292 (2003)  
22889812

JOURNAL

14527350  
2 (bases 1 to 2810)  
Wang, Z.G., Li, L.J., Luo, Y., Zhang, J.Y., Wang, M.Y., Cheng, S.Y.,  
Zhang, Y.J., Wang, X.M., Lu, Y., Wu, N.P., Mei, L.L. and Wang, Z.X.  
Molecular biological analysis of genotyping and phylogeny of severe  
acute respiratory syndrome associated coronavirus  
Chin. Med. J. 117 (1), 42-48 (2004)  
14733771

REFERENCE

3 (bases 1 to 2810)  
Wang, Z., Cheng, S., Zhang, Y., Lu, Y. and Mei, L.  
Direct Submission  
Submitted (03-MAY-2003) Department of Microbiology, Zhejiang Center  
of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou,  
Zhejiang 310009, China  
4 (bases 1 to 2810)  
Wang, Z., Cheng, S., Zhang, Y., Lu, Y. and Mei, L.  
Direct Submission  
Submitted (05-DEC-2003) Department of Microbiology, Zhejiang Center  
of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou,  
Zhejiang 310009, China

REMARK

Nucleotide and amino acid sequences updated by submitter  
On Dec 5, 2003 this sequence version replaced gi:31505969.  
Location/Qualifiers

source

1..2810  
/organism="SARS coronavirus ZJ01"  
/mol\_type="genomic RNA"  
/isolate="ZJ01b"  
/isolation\_source="patient with severe acute respiratory  
syndrome"  
/db\_xref="taxon:230471"  
/country="China: Hangzhou"

misc\_feature

1..6  
/note="transcription regulatory core sequence 6"  
162..353  
/codon\_start=1  
/product="uncharacterized protein 6"  
/protein\_id="AAR27473.1"  
/db\_xref="GI:38735510"  
/translation="MFHLVDQVTAIEIIIMRTFRIAWNLDVLISSIVRQLFKPL  
TKNYSLEDDDEPMELDP"

CDS

331  
/note="Xiaohong genotype locus 6"  
355..360  
/note="transcription regulatory core sequence 7"  
361..729  
/codon\_start=1  
/product="uncharacterized protein 7a"  
/protein\_id="AAR27474.1"  
/db\_xref="GI:38735511"  
/translation="MKILFLTLIVFTSCELYHQCVRGTTVLLKEPCPSCTYBNS  
PFPLADNKKFALTCTSTHFACADGTRHTYQLRARSVPKLFIRQEEVQELYSPLF  
LIVLAALVFLILCFTIKRTE"

misc\_feature

726..860  
/codon\_start=1  
/product="uncharacterized protein 7b"  
/protein\_id="AAR27475.1"  
/db\_xref="GI:38735512"  
/translation="MNEULTIDFYLCFLAFLFLVLIIMLIIFWFSLEIQDLEPCTKV  
"  
861..866  
/note="transcription regulatory core sequence 8"  
867..986

CDS

861..866  
/note="transcription regulatory core sequence 8"  
867..986

misc\_feature

861..866  
/note="transcription regulatory core sequence 8"  
867..986

CDS

861..866  
/note="transcription regulatory core sequence 8"  
867..986

```
/codon_start=1
/product="uncharacterized protein 8a"
/protein_id="AAR27476.1"
/db_xref="GI:38735513"
/translation="MKLLIVLTCSLSCICTVQVRCASKNKPHVLEDPCKVQH"
915 misc_feature
/notes="Xiaohong genotype locus 7"
952..1206
CDS
/codon_start=1
/product="uncharacterized protein 8b"
/protein_id="AAR27477.1"
/db_xref="GI:38735514"
/translation="MCLKILVRYNTRGNTYSTAWLCAIGKVLFPFHRHHTVQVCTPNV
TINCQDPAGALIAACWYIHGEGCTAARFVULVNLKNTIN"
1193..1198
misc_feature
/notes="transcription regulatory core sequence 9"
1208..2476
CDS
/codon_start=1
/product="nucleocapsid protein"
/protein_id="AAP44772.1"
/db_xref="GI:31505970"
/translation="MSDNGPQSNQSRAPRITFGGPTDSTDNNGGRNGARPKQRFPQ
GLPNTATWFTALTQHGKELRPRGQGVPIINTSGPDDQIGYRRATRRVRCGDKGM
KELSPRYFYILCTGEASLPYCANKEGIVWVATEGALNTPKDHICTRNPNNAATVL
QLPQCTTLPGFYAESRGSQASRSRSGNSRNSPTGSRGNSPARMAGGGET
ALALLDLRLNQLSKVSGQGGQGGQGVTKKSAEASAKPKRKRKTATKQYNVTQAFG
RRPEQTQNGFDLIQQTQDYKHFWFOIAQFAPSASFMSRSIRIGMEVTPSGTWLTY
HGAIKLDDKDPQKDNVILINKHIDAYKTFPPTPEPKKKKKKTDQAQPLPQKQKQPT
VTLLPAADMDDFSRQLQNSMSGASADSTQA"
1218..1514
CDS
/codon_start=1
/product="uncharacterized protein 9b"
/protein_id="AAP69659.1"
/db_xref="GI:31747858"
/translation="MDPNQNTVVPVPPALHLVDPOIQTITRMEDAMGQNSADPKVYP
IILRLGQSLSLWARNLDSLEARFQSPFIVQMTKLATTELPDFVVTAK"
1671..1883
CDS
/codon_start=1
/product="uncharacterized protein 9c"
/protein_id="AAP69660.1"
/db_xref="GI:31747859"
/translation="MLPPCYNFKLQHCQKQASTQREAEAAVKELLAPHHVAVIQIQL
LLAAVGEILLEWLAELVKLPSRYCC"

ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 2810;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGAATGAGGACGCAATG 21
|||||
Db 1290 ACCAGAATGAGGACGCAATG 1310

RESULT 9
AY443086S10
LOCUS SARS coronavirus CUHK-L2 orf4 and orf3 genes, partial cds; small
DEFINITION envelope E protein, membrane glycoprotein M, orf7, orf8, orf9,
partial cds; and orf13 gene, complete cds; nucleocapsid protein gene,
partial cds; and orf13 gene, complete cds.
ACCESSION AY443095
VERSION AY443095.1 GI:41352885
KEYWORDS
SEGMENT
SOURCE SARS coronavirus CUHK-L2
ORGANISM SARS coronavirus CUHK-L2
VIRUSES; sRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 3080)
AUTHORS Chim,S.S.C., Tong,Y.K., Hung,E.C.W., Chiu,R.W.K. and Lo,Y.M.D.
TITLE Genomic Sequencing of a SARS Coronavirus Isolate That Predated the
Metropole Hotel Case Cluster in Hong Kong

Clin. Chem. 50 (1), 231-233 (2004)
14709660
2 (bases 1 to 3080)
Chim,S.S.C., Tong,Y.K., Hung,E.C., Chiu,R.W. and Lo,Y.M.
Direct Submission
Submitted (20-OCT-2003) Chemical Pathology, Chinese University of
Hong Kong, Prince of Wales Hospital, 30-32 Ngan Shing Street,
Shatin, Hong Kong, China
Location/Qualifiers
1..3080
/organism="SARS coronavirus CUHK-L2"
/virus
/mol_type="genomic RNA"
/db_xref="taxon:260550"
<1..328
/product="orf2"
/codon_start=2
/product="orf4"
/protein_id="AAS01066.1"
/db_xref="GI:41352897"
/translation="VVILRIGTVLTKMSLYMAISPKFTTSLSHKLQTLVLKMLHS
SLTSLTKTRMCKYQTQSTALQELLIQWIOFMMSRRRLACLCKHKVSTNLCTHSF
RKKQVR"
<1..267
/codon_start=1
/product="orf3"
/protein_id="AAS01065.1"
/db_xref="GI:41352896"
/translation="GGYSEDRHSQKDVVVVHGCVTEVYVQLESTQITTDGTIENATF
FIFNKLIVKPPNVQIHTIDGSSGVANPAMDPIVDEPTTTSVPL"
292..522
CDS
/codon_start=1
/product="small envelope E protein"
/protein_id="AAS01067.1"
/db_xref="GI:41352898"
/translation="MYSFVSEETGLIVNSVLFLAFVVFLLVTLALTALRLCAYCC
NIWVSLVXPTVTVYVYRVKLNSSSEGVPDLIV"
573..1238
CDS
/codon_start=1
/product="membrane glycoprotein M"
/protein_id="AAS01068.1"
/db_xref="GI:41352899"
/translation="MADNGTITVSEELKQLEQNLVIGFLAWIMLLQPAYSNRRPF
LVIIKLVFWLLWPVTLACVLAAYVIRNWTGGIAIACVIGLWMLSYFVASFRFLF
ARTSRMSFNPEPTNILLVPLRGITVITRPMSELVIGAVIIRHLMAGHSLGRCDI
KDLPEITVATSRITLSYKLGASQRVGTDGFAAYNRYRIGNYKLTNDHAGSNDNIAL
LVQ"
1249..1440
CDS
/codon_start=1
/product="orf7"
/protein_id="AAS01069.1"
/db_xref="GI:41352900"
/translation="MFHLVDFQVTVAEIILIIIMRTFRIANWLDVIISIVRQLFKPL
TKKNYSELDDSEPMELDYP"
1448..1816
CDS
/codon_start=1
/product="orf8"
/protein_id="AAS01070.1"
/db_xref="GI:41352901"
/translation="MKIILFLVLTSCELYHQECVRGTTVLLKEPCPSGTYEKNS
PHPLADNRFALTCTSTHFAACADGTRHTYQURARSVSKLFIROEEVQOELYSLPLF
LHVALVFLILCFTTKRTE"
1813..1947
CDS
/codon_start=1
/product="orf9"
/protein_id="AAS01071.1"
/db_xref="GI:41352902"
/translation="MNELTLDIFLCFLAFLFLVLIIMLIIFWFSLEIQDLBEPCTKV
"
1954..2073
CDS
/codon_start=1
/product="orf10"
/protein_id="AAS01072.1"
/db_xref="GI:41352903"
```

```

/translation="MKLLIVLTICSLSCIRTVVQRCSNKPVLIEDPCKVQH"
2039..2293
/codon_start=1
/product="orf11"
/protein_id="AAS01073.1"
/db_xref="GI:41352904"
/translation="MCLKILVYRTRGNTYSTAMLCALGKVLPHRWHTWVQTCPNV
TINCDDPAGGALIAKRWYLGHEGHQTAAPRDLVVLNKRKTN"
2295..>3080
/codon_start=1
/product="nucleocapsid protein"
/protein_id="AAS01074.1"
/db_xref="GI:41352905"
/translation="MSDNGPOSNRSAPRIITFGPTDSTNNQNGRNGARPQKRRPQ
GLPNTASFWTALTOHKEELRFRGQGVPIINTSGPDDQIGYRRATRRVRGGDKM
KELSPWYFYLTGTPEASLPYGANKGIVWVATEXXXXXXXXXXXXXXXXXXXXX
XLPQTTLPKGFYAGSRGSGSOAGSRSSRXSRNSTPGSSRGNSPARWASGGGT
ALALLLDRLAQLESKVSQKQQQGXTVTKSAAEASKPRQK"
2305..2601
/codon_start=1
/product="orf13"
/protein_id="AAS01075.1"
/db_xref="GI:41352906"
/translation="WDPNQTNVVPPALHLVDPQIQLTITRMEDAMGQSONADPKVYP
IILRGSLSLSMARNLDSLEARAFQSTPIVQMTKLATTEELPDEFVVYTK"

ORIGIN
Query Match          100.0%; Score 21; DB 14; Length 3080;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCAGATGGAGGACCAATG 21
|||||
Db 2377 ACCAGATGGAGGACCAATG 2397

RESULT 10
AY304489
LOCUS SARS coronavirus SZ1, partial genome. linear VRL 05-NOV-2003
DEFINITION SARS coronavirus SZ1, partial genome.
ACCESSION AY304489
VERSION AY304489.1 GI:34482140
KEYWORDS
SOURCE SARS coronavirus SZ1
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 8439)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 12958366
PUBMED 12958366
REFERENCE 2 (bases 1 to 8439)
AUTHORS Guan,Y. and Zheng,B.J.
DIRECT SUBMISSION
TITLE Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
Location/Qualifiers
source 1..8439
/organism="SARS coronavirus SZ1"
/mol_type="genomic RNA"
/isolate="SZ1"
/db_xref="taxon:231516"
/country="Hong Kong"

ORIGIN
Query Match          100.0%; Score 21; DB 14; Length 8581;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCAGATGGAGGACCAATG 21
|||||
Db 7066 ACCAGATGGAGGACCAATG 7086

RESULT 12
AY304491
LOCUS SARS coronavirus GZ60, partial genome. linear VRL 05-NOV-2003
DEFINITION SARS coronavirus GZ60, partial genome.
ACCESSION AY304491
VERSION AY304491.1 GI:34482142
KEYWORDS
SOURCE SARS coronavirus GZ60
ORGANISM SARS coronavirus GZ60
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 11006)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 22913660

```

```

Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCAGATGGAGGACCAATG 21
|||||
Db 7072 ACCAGATGGAGGACCAATG 7092

RESULT 11
AY304487
LOCUS SARS coronavirus SZ13, partial genome. linear VRL 05-NOV-2003
DEFINITION SARS coronavirus SZ13, partial genome.
ACCESSION AY304487
VERSION AY304487.1 GI:34482138
KEYWORDS
SOURCE SARS coronavirus SZ13
ORGANISM SARS coronavirus SZ13
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 8581)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 22913660
PUBMED 12958366
REFERENCE 2 (bases 1 to 8581)
AUTHORS Guan,Y. and Zheng,B.J.
DIRECT SUBMISSION
TITLE Submitted (26-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
Location/Qualifiers
source 1..8581
/organism="SARS coronavirus SZ13"
/mol_type="genomic RNA"
/isolate="SZ13"
/db_xref="taxon:231514"
/country="Hong Kong"

ORIGIN
Query Match          100.0%; Score 21; DB 14; Length 8581;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCAGATGGAGGACCAATG 21
|||||
Db 7066 ACCAGATGGAGGACCAATG 7086

RESULT 12
AY304491
LOCUS SARS coronavirus GZ60, partial genome. linear VRL 05-NOV-2003
DEFINITION SARS coronavirus GZ60, partial genome.
ACCESSION AY304491
VERSION AY304491.1 GI:34482142
KEYWORDS
SOURCE SARS coronavirus GZ60
ORGANISM SARS coronavirus GZ60
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 11006)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 22913660

```

```

PUBMED 12958366
REFERENCE 2 (bases 1 to 11006)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..11006
                /organism="SARS coronavirus GZ60"
                /mol_type="genomic RNA"
                /isolate="GZ60"
                /db_xref="taxon:231518"
                /country="Hong Kong"
ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 11006;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCAGAATGGAGCGCAATG 21
    |||||
Db 9473 ACCAGAATGGAGCGCAATG 9493
|||
RESULT 13
AY304493
LOCUS SARS coronavirus HKU-65806 linear VRL 05-NOV-2003
DEFINITION SARS coronavirus HKU-65806, partial genome.
ACCESSION AY304493
VERSION AY304493.1 GI:34482144
KEYWORDS
SOURCE SARS coronavirus HKU-65806
ORGANISM
    Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
    Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 11010)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
PUBMED 12958366
REFERENCE 2 (bases 1 to 11010)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..11010
                /organism="SARS coronavirus HKU-65806"
                /mol_type="genomic RNA"
                /isolate="HKU-65806"
                /db_xref="taxon:231520"
                /country="Hong Kong"
ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 11010;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCAGAATGGAGCGCAATG 21
    |||||
Db 9477 ACCAGAATGGAGCGCAATG 9497
|||
RESULT 14
AY304494
LOCUS SARS coronavirus GZ43 linear VRL 05-NOV-2003
DEFINITION SARS coronavirus GZ43, partial genome.
ACCESSION AY304490
VERSION AY304490.1 GI:34482141
KEYWORDS
SOURCE SARS coronavirus GZ43
ORGANISM
    Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
    Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 13471)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
PUBMED 12958366
REFERENCE 2 (bases 1 to 13471)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..13471
                /organism="SARS coronavirus GZ43"
                /mol_type="genomic RNA"

```

```

DEFINITION SARS coronavirus HKU-66078, partial genome.
ACCESSION AY304494
VERSION AY304494.1 GI:34482145
KEYWORDS
SOURCE SARS coronavirus HKU-66078
ORGANISM
    Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
    Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 11010)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
PUBMED 12958366
REFERENCE 2 (bases 1 to 11010)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..11010
                /organism="SARS coronavirus HKU-66078"
                /mol_type="genomic RNA"
                /isolate="HKU-66078"
                /db_xref="taxon:231521"
                /country="Hong Kong"
ORIGIN
Query Match 100.0%; Score 21; DB 14; Length 11010;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCAGAATGGAGCGCAATG 21
    |||||
Db 9477 ACCAGAATGGAGCGCAATG 9497
|||
RESULT 15
AY304490
LOCUS SARS coronavirus GZ43 linear VRL 05-NOV-2003
DEFINITION SARS coronavirus GZ43, partial genome.
ACCESSION AY304490
VERSION AY304490.1 GI:34482141
KEYWORDS
SOURCE SARS coronavirus GZ43
ORGANISM
    Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
    Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 13471)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
PUBMED 12958366
REFERENCE 2 (bases 1 to 13471)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..13471
                /organism="SARS coronavirus GZ43"
                /mol_type="genomic RNA"

```

/isolate="GZ43"  
/db\_xref="taxon:231517"  
/country="Hong Kong"

ORIGIN

Query Match 100.0%; Score 21; DB 14; Length 13471;  
Best Local Similarity 100.0%; Pred. No. 7.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCAGATGGAGGACGCAATG 21  
|||  
Db 11938 ACCAGATGGAGGACGCAATG 11958

Search completed: May 16, 2005, 02:00:25  
Job time : 503.031 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 17:17:00 ; Search time 269.391 Seconds  
(without alignments)  
461.466 Million cell updates/sec

Title: US-10-808-187A-2474

Perfect score: 21

Sequence: 1 accagaatgagacgcaatg 21

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	29751	12	ADJ39000 SARS coro
2	17.8	84.8	4457	4	ABL06791 Drosophil
3	17.8	84.8	9051	4	ABL06790 Drosophil
4	16.8	80.0	1250	2	AAQ50571 Asparagin
5	16.8	80.0	1453	2	AAQ50573 Asparagin
6	16.8	80.0	1640	2	AAQ50575 Asparagin
7	16.8	80.0	1910	2	AAQ50579 Asparagin
8	16.8	80.0	10419	3	AAAS5188 Cenarchae
9	16.8	80.0	27082	4	AAK70447 Human imm
10	16.8	80.0	42432	3	AAAS5187
11	16.8	80.0	179651	10	ADL13813 Osteoarth
12	16.4	78.1	646	4	AAH83969 Papio ham
13	16.4	78.1	646	4	AAH83967 Papio ham
14	16.4	78.1	63411	12	ADQ97081 Mouse can
15	16.2	77.1	192	4	AAQ62139 Murine be
16	16.2	77.1	448	4	ABA58687 Human foe
17	16.2	77.1	448	4	AAI38366 Probe #70
18	16.2	77.1	448	4	AAK32540 Human bon
19	16.2	77.1	448	4	AAK06821 Human bra
20	16.2	77.1	448	4	ABS32250 Human liv

21	16.2	77.1	448	6	ABS07329 Human gen
22	16.2	77.1	1092	6	ABK72777 Bacillus
23	16.2	77.1	1227	8	ACA51603 Prokaryot
24	16.2	77.1	1248	11	ACH96177 Klebsiell
25	16.2	77.1	1251	8	ACA48608 Prokaryot
26	16.2	77.1	3228	5	AS81183 DNA encod
27	16.2	77.1	7744	4	AAF83380 P. chryso
28	16.2	77.1	7744	4	AAF83381 P. chryso
29	16.2	77.1	87464	11	ACN44788 Mouse gen
30	16.2	77.1	349980	5	AAH41223 Pyrococcu
31	16.2	77.1	349980	6	ABQ81846 Bifidobac
32	16	76.2	1849	10	ADI62671 Human apo
33	16	76.2	1926	6	ABN84482 Rice AINI
34	16	76.2	7817	13	ADR84544 Aspergill
35	15.8	75.2	258	3	AAA82225 N. mening
36	15.8	75.2	382	5	AAF67493 Novel hum
37	15.8	75.2	443	6	ABL94092 Arabidops
38	15.8	75.2	452	2	AAV68987 DNA molec
39	15.8	75.2	452	3	AAQ80990 Human bre
40	15.8	75.2	452	6	AAAS99836 Breast tu
41	15.8	75.2	452	6	ABK46880 Human bre
42	15.8	75.2	452	8	ADAI1357 Human bre
43	15.8	75.2	452	10	ADC15330 Human bre
44	15.8	75.2	481	2	AAQ29870 Pheromone
45	15.8	75.2	481	10	ADI04956 Rat Odora

## ALIGNMENTS

### RESULT 1

ADJ39000

ID ADJ39000 standard; DNA; 29751 BP.

XX AC ADJ39000;

XX DT 06-MAY-2004 (first entry)

XX DE SARS coronavirus nucleotide sequence.

XX KW small interfering RNA; siRNA; modified ribonucleotide;

XX KW viral replication inhibition; hepatitis C virus; HCV; hepatitis C;

XX KW antiinflammatory; hepatotropic; virucide; hepatitis A virus;

XX KW hepatitis D virus; hepatitis E virus; Ebola virus; influenza virus;

XX KW rotavirus; reovirus; retrovirus; poliovirus; human papilloma virus;

XX KW metapneumoniavirus; coronavirus; viral infection; gene; ds.

XX OS SARS coronavirus.

XX WO2004011647-A1.

XX PD 05-FEB-2004.

XX PF 25-JUL-2003; 2003WO-US023104.

XX PR 26-JUL-2002; 2002US-0398605P.

XX PA (CHIR ) CHIRON CORP.

XX PI Han J, Seo MY, Houghton M;

XX DR WPI; 2004-143862/14.

XX PT New RNase resistant small interfering RNA, useful for treating viral

XX PT infections, e.g., hepatitis C, influenza virus or coronavirus infection.

XX PS Example 10; Fig 3; 74pp; English.

XX CC The present invention describes a small interfering RNA (siRNA) which

XX CC comprises a modified ribonucleotide, where the siRNA is resistant to

XX CC RNase and retains the ability to inhibit viral replication. Also

XX CC described: (1) inactivating a virus in a patient; (2) making a modified

XX CC siRNA that targets a nucleic acid sequence in a virus; (3) a double-

CC stranded RNA molecule of 10-30 nucleotides that inhibits replication of  
 CC hepatitis C virus (HCV); (4) inducing targeted RNA interference toward  
 CC HCV in hepatic cells; (5) inhibiting replication of HCV; (6) a vector  
 CC comprising a DNA segment encoding the RNA molecule; (7) a host cell  
 CC comprising the vector of (6); (8) inhibiting replication of HCV in cells  
 CC carrying HCV; (9) treating hepatitis C in a subject; (10) a modified  
 CC siRNA molecule comprising a double-stranded RNA molecule of 10-30  
 CC nucleotides in length, which mediates RNA interference toward a target  
 CC agent or virus and is linked to at least one receptor-binding ligand; and  
 CC (11) inducing targeted RNA interference in a patient. The modified siRNA  
 CC molecules have antiinflammatory, hepatotropic and virucide activities.  
 CC The modified RNA molecules are useful for inactivating virus in mammalian  
 CC cells. The siRNAs are useful for treating hepatitis C virus, hepatitis A  
 CC virus, hepatitis D virus, hepatitis E virus, Ebola virus, influenza  
 CC virus, rotavirus, reovirus, retrovirus, poliovirus, human papilloma  
 CC virus, metapneumoniavirus or coronavirus infections. The methods of the  
 CC invention can be used to correct or compensate for cellular physiological  
 CC abnormalities involved in conferring susceptibility to viral infections  
 CC in patients and/or alleviate symptoms of a viral infection in patients.  
 CC The present sequence represents the SARS coronavirus nucleotide sequence,  
 CC which is used in an example from the present invention.

SQ Sequence 29751 BP; 8478 A; 5941 C; 6188 G; 9144 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 12; Length 29751;  
 Best Local Similarity 100.0%; Pred. No. 3.9;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGACGCAATG 21

|||||||

Db 28202 ACCAGAATGGAGGACGCAATG 28222

RESULT 2

ABL06791

ID ABL06791 standard; cDNA; 4457 BP.

AC ABL06791;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 14855.

CC Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-PSDB; ABB62687.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 CC genes from Drosophila and for elucidating cell signaling and cell-cell  
 CC interactions.

PS Claim 1; SEQ ID NO 14855; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB161176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 4457 BP; 1222 A; 1176 C; 1212 G; 847 T; 0 U; 0 Other;

Query Match 84.8%; Score 17.8; DB 4; Length 4457;  
 Best Local Similarity 90.5%; Pred. No. 1.2e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGACGCAATG 21

|||||||

Db 2592 ACCAGAATGGAGGACGCAATG 2612

RESULT 3

ABL06790/c

ID ABL06790 standard; cDNA; 9051 BP.

XX ABL06790;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 14852.

CC Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-PSDB; ABB62687.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 CC genes from Drosophila and for elucidating cell signaling and cell-cell  
 CC interactions.

PS Claim 1; SEQ ID NO 14852; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB161176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 9051 BP; 2144 A; 2170 C; 2117 G; 2620 T; 0 U; 0 Other;

Query Match 84.8%; Score 17.8; DB 4; Length 9051;

Best Local Similarity 90.5%; Pred. No. 1.3e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGAGCGCAATG 21  
 Db 2942 ACCAGATGGAGGAGCGAGTG 2922

## RESULT 4

AAQ50571/c  
 ID AAQ50571 standard; cDNA to mRNA; 1250 BP.

XX AC AAQ50571;  
 XX 25-MAR-2003 (revised)  
 DT 24-MAY-1994 (first entry)  
 XX

DE Asparaginylendopeptidase clone 102.

XX Asparaginylendopeptidase; Canavalia ensiformis; seed; L-asparagine;  
 KW primer; PCR; protein fragmentation; peptide synthesis; ss.

XX Canavalia ensiformis.

XX Key Location/Qualifiers  
 FH CDS 1..957  
 FT /\*tag= a  
 XX

PN JP05276960-A.

XX 26-OCT-1993.

XX 07-AUG-1992; 92JP-00231602.

XX 07-FEB-1992; 92JP-00056023.

XX (TAKI ) TAKARA SHUZO CO LTD.

XX WPI; 1993-373587/47.

DR P-PSDB; AAR43036.

XX New gene for encoding asparaginyl endo-peptidase - comprises 8 specified  
 PT DNA sequences.

XX Disclosure; Page 19-20; 35pp; Japanese.

XX A gene encoding asparaginylendopeptidase is claimed. 8 DNA sequences are  
 CC given (AAQ50559-66). The enzyme is a protease derived from a seed of  
 CC Canavalia ensiformis which selectively hydrolyses C-terminus amide bond  
 CC of L-asparagine residue (see AAR43033 and AAR43041). The enzyme is useful  
 CC for protein fragmentation and enzymatic peptide synthesis. The primers  
 CC given in AAQ50567-68, AAQ50576-77 and AAQ50583-90 were used in the  
 CC isolation of the fragments given in AAQ50569-75 and AAQ50578-79, by PCR.  
 CC (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 1250 BP; 370 A; 229 C; 298 G; 353 T; 0 U; 0 Other;

Query Match 80.0%; Score 16.8; DB 2; Length 1250;  
 Best Local Similarity 90.0%; Pred. No. 3.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGAGCGCAAT 20  
 Db 1049 ACCAGAAGGGAGGAACGCAAT 1030

## RESULT 5

AAQ50573/c  
 ID AAQ50573 standard; cDNA to mRNA; 1453 BP.

XX AC AAQ50573;  
 XX 25-MAR-2003 (revised)  
 DT 24-MAY-1994 (first entry)  
 XX

DE Asparaginylendopeptidase clone 104.

XX Asparaginylendopeptidase; Canavalia ensiformis; seed; L-asparagine;  
 KW primer; PCR; protein fragmentation; peptide synthesis; ss.

XX Canavalia ensiformis.

XX Key Location/Qualifiers  
 FH CDS 3..1094  
 FT /\*tag= a  
 XX

PN JP05276960-A.

XX 26-OCT-1993.

XX 07-AUG-1992; 92JP-00231602.

XX 07-FEB-1992; 92JP-00056023.

XX (TAKI ) TAKARA SHUZO CO LTD.

XX WPI; 1993-373587/47.

DR P-PSDB; AAR43038.

XX New gene for encoding asparaginyl endo-peptidase - comprises 8 specified  
 PT DNA sequences.

XX Disclosure; Page 22-24; 35pp; Japanese.

XX A gene encoding asparaginylendopeptidase is claimed. 8 DNA sequences are  
 CC given (AAQ50559-66). The enzyme is a protease derived from a seed of  
 CC Canavalia ensiformis which selectively hydrolyses C-terminus amide bond  
 CC of L-asparagine residue (see AAR43033 and AAR43041). The enzyme is useful  
 CC for protein fragmentation and enzymatic peptide synthesis. The primers  
 CC given in AAQ50567-68, AAQ50576-77 and AAQ50583-90 were used in the  
 CC isolation of the fragments given in AAQ50569-75 and AAQ50578-79, by PCR.  
 CC (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 1453 BP; 480 A; 252 C; 335 G; 386 T; 0 U; 0 Other;

Query Match 80.0%; Score 16.8; DB 2; Length 1453;  
 Best Local Similarity 90.0%; Pred. No. 3.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGAGCGCAAT 20  
 Db 1186 ACCAGAAGGGAGGAACGCAAT 1167

## RESULT 6

AAQ50575/c

ID AAQ50575 standard; cDNA to mRNA; 1640 BP.

XX AC AAQ50575;

XX 25-MAR-2003 (revised)

DT 24-MAY-1994 (first entry)

DE Asparaginylendopeptidase clone ASN-1.

XX Asparaginylendopeptidase; Canavalia ensiformis; seed; L-asparagine;  
 KW primer; PCR; protein fragmentation; peptide synthesis; ss.

XX Canavalia ensiformis.

PN JP05276960-A.

XX 26-OCT-1993.

XX 07-AUG-1992; 92JP-00231602.

XX 07-FEB-1992; 92JP-00056023.

XX (TAKI ) TAKARA SHUZO CO LTD.

```

XX WPI; 1993-373587/47.
XX
XX New gene for encoding asparaginyl endo-peptidase - comprises 8 specified
XX DNA sequences.
XX
XX Disclosure; Page 26; 35pp; Japanese.
XX
XX A gene encoding asparaginylendopeptidase is claimed. 8 DNA sequences are
XX given (AAQ50559-66). The enzyme is a protease derived from a seed of
XX Canavalia ensiformis which selectively hydrolyses C-terminus amide bond
XX of L-asparagine residue (see AAR43033 and AAR43041). The enzyme is useful
XX for protein fragmentation and enzymatic peptide synthesis. The primers
XX given in AAQ50567-68, AAQ50576-77 and AAQ50583-90 were used in the
XX isolation of the fragments given in AAQ50569-75 and AAQ50578-79, by PCR.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 1640 BP; 533 A; 288 C; 383 G; 436 T; 0 U; 0 Other;
XX
XX Query Match 80.0%; Score 16.8; DB 2; Length 1640;
XX Best Local Similarity 90.0%; Pred. No. 3.2e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 ACCAGAAATGGAGGACGCAAT 20
Db 1373 ACCAGAAAGGGAGACGCAAT 1354

RESULT 7
AAQ50579/c
ID AAQ50579 standard; cDNA to mRNA; 1910 BP.
XX
XX AAQ50579;
XX
XX 25-MAR-2003 (revised)
XX 24-MAY-1994 (first entry)
XX
XX Asparaginylendopeptidase ASN.
XX
XX Asparaginylendopeptidase; Canavalia ensiformis; seed; L-asparagine;
XX primer; PCR; protein fragmentation; peptide synthesis; ss.
XX
XX Canavalia ensiformis.
XX
XX Key Location/Qualifiers
XX CDS 229..1551
XX FT /*tag= a
XX
XX JP05276960-A.
XX
XX 26-OCT-1993.
XX
XX 07-AUG-1992; 92JP-00231602.
XX
XX 07-FEB-1992; 92JP-00056023.
XX
XX (TAKI ) TAKARA SHUZO CO LTD.
XX
XX WPI; 1993-373587/47.
XX P-PSDB; AAR43040.
XX
XX New gene for encoding asparaginyl endo-peptidase - comprises 8 specified
XX DNA sequences.
XX
XX Disclosure; Page 27-29; 35pp; Japanese.
XX
XX A gene encoding asparaginylendopeptidase is claimed. 8 DNA sequences are
XX given (AAQ50559-66). The enzyme is a protease derived from a seed of
XX Canavalia ensiformis which selectively hydrolyses C-terminus amide bond
XX of L-asparagine residue (see AAR43033 and AAR43041). The enzyme is useful
XX for protein fragmentation and enzymatic peptide synthesis. The primers
XX given in AAQ50567-68, AAQ50576-77 and AAQ50583-90 were used in the
XX isolation of the fragments given in AAQ50569-75 and AAQ50578-79, by PCR.
XX

```

---

```

CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 1910 BP; 591 A; 339 C; 464 G; 516 T; 0 U; 0 Other;
XX
XX Query Match 80.0%; Score 16.8; DB 2; Length 1910;
XX Best Local Similarity 90.0%; Pred. No. 3.2e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 ACCAGAAATGGAGGACGCAAT 20
Db 1643 ACCAGAAAGGGAGACGCAAT 1624

RESULT 8
AAAS5188
ID AAAS5188 standard; DNA; 10419 BP.
XX
XX AAAS5188;
XX
XX 30-AUG-2000 (first entry)
XX
XX Cenarchaeum symbiosum open reading frame nucleotide sequence SEQ ID NO:3.
XX
XX Cenarchaeum symbiosum; non-thermophilic; crenarchaeote; physiology;
XX characterisation; archae; therapeutic; industrial; laboratory; ds.
XX
XX Cenarchaeum symbiosum.
XX
XX WO200018909-A2.
XX
XX 06-APR-2000.
XX
XX 29-SEP-1999; 99WO-US022752.
XX
XX 29-SEP-1998; 98US-0102294P.
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Swanson RV, Feldman RA, Schleper C;
XX
XX WPI; 2000-293148/25.
XX P-PSDB; AAY90913.
XX
XX New nucleic acids and proteins isolated from the non-thermophilic
XX crenarchaeote Cenarchaeum symbiosum, useful in characterizing the
XX physiology of these archae and in therapeutic, industrial or laboratory
XX techniques.
XX
XX Claim 15; Page 87-102; 210pp; English.
XX
XX AAAS5186 to AAAS5226 and AAY90913 to AAY90951 represent nucleic acids and
XX proteins isolated from the non-thermophilic crenarchaeote Cenarchaeum
XX symbiosum. The nucleic acids and proteins identified in the present
XX invention are useful in characterising the physiology of these archae and
XX can be used in therapeutic, industrial or laboratory techniques. AAAS5227
XX to AAAS5260 represent promoter sequences from Cenarchaeum symbiosum.
XX AAAS5261 to AAAS5269 represent PCR primers and probes used in examples
XX from the present invention
XX
XX Sequence 10419 BP; 2064 A; 2958 C; 3341 G; 2056 T; 0 U; 0 Other;
XX
XX Query Match 80.0%; Score 16.8; DB 3; Length 10419;
XX Best Local Similarity 90.0%; Pred. No. 4e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 2 CCAGAAATGGAGGACGCAATG 21
Db 9044 CCAGAGTGGAGGATGCAATG 9063

RESULT 9
AAK70447
ID AAK70447 standard; DNA; 27082 BP.

```

XX AC AAK70447;  
XX DT 06-NOV-2001 (first entry)  
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25259.  
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cyostatic; gene therapy; vaccine; metastasis; ds.  
XX OS Homo sapiens.  
XX PN WO200157182-A2.  
XX PD 09-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US001354.  
XX PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249255P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-02556719P.

```

PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 25259; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 27082 BP; 5963 A; 7786 C; 7711 G; 5622 T; 0 U; 0 Other;
Query Match 80.0%; Score 16.8; DB 4; Length 27082;
Best Local Similarity 90.0%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCAGATGGAGGACGCAATG 21
Db 24699 CCAGATGGAGGACGCACTG 24718
RESULT 10
AAAS5187
ID AAAS5187 standard; DNA; 42432 BP.
AC
XX AAAS5187;
XX
XX 30-AUG-2000 (first entry)
XX
XX Cenarchaeum symbiosum nucleotide sequence variant B SEQ ID NO:2.
XX
XX Cenarchaeum symbiosum; non-thermophilic; crenarchaeote; physiology;
XX characterisation; archae; therapeutic; industrial; laboratory; ds.
XX
XX Cenarchaeum symbiosum.
OS
XX WO200018909-A2.
XX
XX 06-APR-2000.
XX
XX 29-SEP-1999; 99WO-US022752.
XX
XX 29-SEP-1998; 98US-0102294P.
XX
XX (DIVE-) DIVERSA CORP.
PA

```

```

XX Swanson RV, Feldman RA, Schleper C;
XX WPI; 2000-293148/25.
XX
XX New nucleic acids and proteins isolated from the non-thermophilic
XX crenarchaeote Cenarchaeum symbiosum, useful in characterizing the
XX physiology of these archae and in therapeutic, industrial or laboratory
XX techniques.
XX
XX Claim 1; Page 75-87; 210pp; English.
XX
XX AAAS5186 to AAAS5226 and AAY90913 to AAY90951 represent nucleic acids and
XX proteins isolated from the non-thermophilic crenarchaeote Cenarchaeum
XX symbiosum. The nucleic acids and proteins identified in the present
XX invention are useful in characterising the physiology of these archae and
XX can be used in therapeutic, industrial or laboratory techniques. AAAS5227
XX to AAAS5260 represent promoter sequences from Cenarchaeum symbiosum.
XX AAAS5261 to AAAS5269 represent PCR primers and probes used in examples
XX from the present invention
XX
XX Sequence 42432 BP; 8792 A; 12248 C; 12606 G; 8786 T; 0 U; 0 Other;
Query Match 80.0%; Score 16.8; DB 3; Length 42432;
Best Local Similarity 90.0%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCAGATGGAGGACGCAATG 21
Db 9046 CCAGATGGAGGACGCAATG 9065
RESULT 11
ADL13813/c
ID ADL13813 standard; DNA; 179651 BP.
AC ADL13813;
XX
XX 06-MAY-2004 (first entry)
XX
XX Osteoarthritis-associated polymorphic nucleotide #345.
XX
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX joint space narrowing; osteophyte development; joint pain;
XX osteoarthritis; SNP; single nucleotide polymorphism.
XX
XX Homo sapiens.
OS
XX WO2003054166-A2.
XX
XX 03-JUL-2003.
XX
XX 19-DEC-2002; 2002WO-US041225.
XX
XX 20-DEC-2001; 2001US-0342603P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Jones KA, Schafer A;
XX WPI; 2003-559141/52.
XX
XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.
XX
XX Disclosure; SEQ ID NO 345; 297pp; English.
XX
XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein

```

CC listed in the specification. The methods, composition and agent are  
 CC useful for modulating the susceptibility of an individual to joint space  
 CC narrowing and/or osteophyte development and/or joint pain that is  
 CC associated with a disease, preferably osteoarthritis. The cell line and  
 CC the non-human animal are useful for screening for an agent for diagnosing  
 CC an individual having susceptibility to joint space narrowing and/or  
 CC osteophyte development and/or joint pain. This sequence corresponds to  
 CC the polynucleotide encoding a protein listed in the specification. (Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences).

XX Sequence 179651 BP; 48335 A; 39400 C; 42185 G; 49731 T; 0 U; 0 Other;

Query Match 80.0%; Score 16.8; DB 10; Length 179651;  
 Best Local Similarity 90.0%; Pred. No. 5.7e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCAGATGGAGGAGCGCAATG 21

DB 157352 CAAGATGGAGGAGCAATG 157333

RESULT 12

AAH83969/c

ID AAH83969 standard; DNA; 646 BP.

XX AC AAH83969;

DT 25-SEP-2001 (first entry)

DE Papio hamadryas olfactory receptor pseudogene 3.

XX Olfactory receptor; primate; mouse; human; food processing industry;  
 KW aromas; perfumery; toxic substance; ds.

XX Papio hamadryas.

XX WO200146262-A2.

XX 28-JUN-2001.

XX 22-DEC-2000; 2000WO-IB002017.

XX 22-DEC-1999; 99US-0171746P.

PR 21-DEC-2000; 2000US-00747155.

XX (CNRS ) CNRS CENT NAT RECH SCI.

PI Rouquier S, Giorgi D;

XX WPI; 2001-381911/40.

XX Nucleic acids encoding primate and murine olfactory receptors, useful for  
 PT analysis odors e.g. in food processing and perfumery.

XX Claim 1; Page 195-196; 482pp; English.

CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the  
 CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10  
 CC primate species, mouse and human. The nucleic acids and receptors may be  
 CC used in the food processing industry (e.g. for the detection of aromas,  
 CC quality control and sample analysis), in perfumery (e.g. for the analysis  
 CC or comparison of perfumes) and in the environment (e.g. for the detection  
 CC of toxic substances and/or trapping of odours)

XX Sequence 646 BP; 136 A; 167 C; 122 G; 221 T; 0 U; 0 Other;

Query Match 78.1%; Score 16.4; DB 4; Length 646;

Best Local Similarity 94.4%; Pred. No. 4.4e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CAGAATGGAGGAGCGCAAT 20

DB 478 CAGAATGGAGGAGCGCAAT 461

RESULT 13

AAH83967/c

ID AAH83967 standard; DNA; 646 BP.

XX AC AAH83967;

DT 25-SEP-2001 (first entry)

DE Papio hamadryas olfactory receptor encoding gene 10.

XX Olfactory receptor; primate; mouse; human; food processing industry;  
 KW aromas; perfumery; toxic substance; ds.

XX Papio hamadryas.

XX WO200146262-A2.

XX 28-JUN-2001.

XX 22-DEC-2000; 2000WO-IB002017.

XX 22-DEC-1999; 99US-0171746P.

PR 21-DEC-2000; 2000US-00747155.

XX (CNRS ) CNRS CENT NAT RECH SCI.

PI Rouquier S, Giorgi D;

XX WPI; 2001-381911/40.

DR P-PSDB; AAG98485.

XX Nucleic acids encoding primate and murine olfactory receptors, useful for  
 PT analysis odors e.g. in food processing and perfumery.

XX Claim 1; Page 193; 482pp; English.

CC The invention relates to olfactory receptors (AAG98432-AAG98609) and the  
 CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10  
 CC primate species, mouse and human. The nucleic acids and receptors may be  
 CC used in the food processing industry (e.g. for the detection of aromas,  
 CC quality control and sample analysis), in perfumery (e.g. for the analysis  
 CC or comparison of perfumes) and in the environment (e.g. for the detection  
 CC of toxic substances and/or trapping of odours)

XX Sequence 646 BP; 135 A; 166 C; 123 G; 222 T; 0 U; 0 Other;

Query Match 78.1%; Score 16.4; DB 4; Length 646;

Best Local Similarity 94.4%; Pred. No. 4.4e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CAGAATGGAGGAGCGCAAT 20

DB 478 CAGAATGGAGGAGCGCAAT 461

RESULT 14

ADQ97081/c

ID ADQ97081 standard; DNA; 63411 BP.

XX AC - ADQ97081;

DT 07-OCT-2004 (first entry)

DE Mouse cancer associated sequence MD10-004, SEQ ID 57.

XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.

XX Mus musculus.

```
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
XX 22-DEC-2003; 2003WO-US041389.
XX
XX 27-DEC-2002; 2002US-00330773.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
XX
XX WPI; 2004-543781/52.
XX
XX New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
XX Claim 1; SEQ ID NO 57; 199pp; English.
XX
XX The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 63411 BP; 17408 A; 15472 C; 14344 G; 15847 T; 0 U; 340 Other;
SQ
Query Match 78.1%; Score 16.4; DB 12; Length 63411;
Best Local Similarity 94.4%; Pred. No. 7.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 AGAATGGAGGACGCAATG 21
DB 18892 AGAATGGAGGACGCAATG 18875
RESULT 15
AAC62139/c
ID AAC62139 standard; DNA; 192 BP.
XX
XX AAC62139;
XX
XX 06-MAR-2001 (first entry)
XX
XX Murine beta-1-6-N-acetylglucosaminyltransferase DNA fragment.
XX
XX Human; beta-1-6-N-acetylglucosaminyltransferase; C2GnT-M; inflammation;
KW membrane protein; branched sialyl Lex; L-selectin; immune reaction;
KW inflammation; tissue rejection; tumour metastasis; ss.
XX
XX Mus musculus.
XX
XX US6136580-A.
XX
XX 24-OCT-2000.
XX
XX 19-JAN-1999; 99US-00233506.
XX
XX 19-JAN-1999; 99US-00233506.
XX
XX (BURN-) BURNHAM INST.
XX
XX Fukuda M, Yeh J;
XX
XX WPI; 2001-040238/05.
XX
XX P-PSDB; AAE30519.
XX
XX New C2GnT-M polypeptides having core 2, core 4 and I branching beta-1-6-N
PT -acetylglucosaminyltransferase activities for preparing reagents useful
PT for diagnosing, preventing or treating inflammation or tumor metastasis.
XX
```

```
PS Example 5; Col 33-34; 25pp; English.
XX
XX The present sequence encodes a beta-1-6-N-acetylglucosaminyltransferase
CC fragment. The full length polypeptide has core2, core4 and I branching
CC activities. It is designated C2GnT-M. C2GnT-M is a membrane protein that
CC is predominantly expressed in colon, small intestine, trachea, stomach
CC and thyroid, as well as in certain cancer cell lines. C2GnT-M
CC polypeptides may be used to prepare molecules having highly branched
CC sialyl Lex and L-selectins, which may be subsequently used to modulate
CC immune reactions, e.g. inflammation and tissue rejection, and to prevent
XX or inhibit tumour metastasis
XX
XX Sequence 192 BP; 48 A; 48 C; 51 G; 45 T; 0 U; 0 Other;
SQ
Query Match 77.1%; Score 16.2; DB 4; Length 192;
Best Local Similarity 85.7%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCAGATGGAGGACGCAATG 21
DB 59 ACCACCATGGAGTACGCAATG 39
Search completed: May 16, 2005, 01:09:31
Job time : 274.391 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:11:05 ; Search time 78.5859 Seconds  
(without alignments)  
437.251 Million cell updates/sec

Title: US-10-808-187A-2474

Perfect score: 21

Sequence: 1 accagatggaggacgaatg 21

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS-COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	80.0	10419	4	US-09-408-020-3
2	16.8	80.0	42432	4	US-09-408-020-2
3	16.8	80.0	75799	4	US-09-949-016-15231
4	16.2	77.1	192	3	US-09-233-506-9
5	16.2	77.1	1248	4	US-09-489-039A-1972
6	16.2	77.1	56976	4	US-09-949-016-17486
7	16.2	77.1	101674	4	US-09-949-016-12033
8	15.8	75.2	452	3	US-08-991-789A-278
9	15.8	75.2	452	3	US-09-062-451-278
10	15.8	75.2	452	4	US-09-289-198-278
11	15.8	75.2	452	4	US-09-429-755-278
12	15.8	75.2	452	4	US-09-699-295-278
13	15.8	75.2	601	4	US-09-949-016-81705
14	15.8	75.2	601	4	US-09-949-016-81706
15	15.8	75.2	601	4	US-09-949-016-81707
16	15.8	75.2	601	4	US-09-949-016-174315
17	15.8	75.2	601	4	US-09-949-016-174507
18	15.8	75.2	1701	3	US-09-232-468A-21
19	15.8	75.2	1701	4	US-09-784-984B-17
20	15.8	75.2	1777	1	US-08-229-781-54
21	15.8	75.2	1777	1	US-08-630-918-54
22	15.8	75.2	1777	3	US-09-004-422-54
23	15.8	75.2	1777	4	US-09-918-568-54
24	15.8	75.2	125192	4	US-09-949-016-14120
25	15.8	75.2	126237	4	US-09-949-016-16674
26	15.8	75.2	126237	4	US-09-949-016-16675
27	15.8	75.2	152331	3	US-09-128-155-16

Sequence 3746, Ap  
Sequence 19028, A  
Sequence 15, Appli  
Sequence 1, Appli  
Sequence 14184, A  
Sequence 5346, Ap  
Sequence 3362, Ap  
Sequence 27937, A  
Sequence 128820,  
Sequence 1359, Ap  
Sequence 2602, Ap  
Sequence 13, Appl  
Sequence 17, Appl  
Sequence 20, Appl  
Sequence 5, Appli  
Sequence 8, Appli  
Sequence 5, Appli  
Sequence 5, Appli

US-09-270-767-3746  
US-09-270-767-19028  
US-09-936-145-15  
US-09-990-337-1  
US-09-949-016-14184  
US-09-252-991A-5346  
US-09-248-796A-3362  
US-09-949-016-27937  
US-09-949-016-128820  
US-09-583-110-1359  
US-09-107-433-2602  
US-08-753-247-13  
US-08-753-247-17  
US-08-753-247-20  
US-08-753-247-5  
US-08-753-247-8  
US-08-477-254A-5  
US-08-472-576B-5

## ALIGNMENTS

### RESULT 1

US-09-408-020-3

; Sequence 3, Application US/09408020

; Patent No. 6632937

; GENERAL INFORMATION:

; APPLICANT: Swanson, Ronald V.

; APPLICANT: Feldman, Robert A.

; APPLICANT: Schleper, Christa

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM

; FILE REFERENCE: DCOF.002A

; CURRENT APPLICATION NUMBER: US/09/408,020

; CURRENT FILING DATE: 1999-09-29

; PRIOR APPLICATION NUMBER: 60/102,294

; PRIOR FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 123

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 10419

; TYPE: DNA

; ORGANISM: Cenarchaeum symbiosum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(10419)

US-09-408-020-3

Query Match

Best Local Similarity 80.0%; Score 16.8; DB 4; Length 10419;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCAGATGGAGGACGAATG 21

Db 9044 CCAGATGGAGGACGAATG 9063

### RESULT 2

US-09-408-020-2

; Sequence 2, Application US/09408020

; Patent No. 6632937

; GENERAL INFORMATION:

; APPLICANT: Swanson, Ronald V.

; APPLICANT: Feldman, Robert A.

; APPLICANT: Schleper, Christa

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM

; FILE REFERENCE: DCOF.002A

; CURRENT APPLICATION NUMBER: US/09/408,020

; CURRENT FILING DATE: 1999-09-29

; PRIOR APPLICATION NUMBER: 60/102,294

; PRIOR FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 123

; SOFTWARE: FastSeq for Windows Version 3.0

```
; SEQ ID NO 2
; LENGTH: 42432
; TYPE: DNA
; ORGANISM: Cenarchaeum symbiosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3) ... (10421)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10625) ... (11434)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11478) ... (13046)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13046) ... (14620)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23558) ... (24862)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24913) ... (25728)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26504) ... (26881)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (29655) ... (30491)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34559) ... (36067)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37002) ... (37403)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37404) ... (38282)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39454) ... (40572)
; US-09-408-020-2
;
; Query Match
; Best Local Similarity 80.0%; Score 16.8; DB 4; Length 42432;
; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 2 CCAGATGGAGGACGCAATG 21
; Db 9046 CCAGATGGAGGATGCAATG 9065
;
; RESULT 3
; US-09-949-016-15231/c
; Sequence 15231, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15231
; LENGTH: 75799
; TYPE: DNA
```

```
; ORGANISM: Human
; US-09-949-016-15231
;
; Query Match
; Best Local Similarity 80.0%; Score 16.8; DB 4; Length 75799;
; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 2 CCAGATGGAGGACGCAATG 21
; Db 43216 CCAGATGGAGGACCAATG 43197
;
; RESULT 4
; US-09-233-506-9/c
; Sequence 9, Application US/09233506
; Patent No. 6136580
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Minoru
; APPLICANT: Yeh, Jiunn-Chern
; TITLE OF INVENTION: A Beta-1-6-N-Acetylglucosaminyltransferase That Forms
; FILE REFERENCE: P-LJ 3415
; CURRENT APPLICATION NUMBER: US/09/233,506
; CURRENT FILING DATE: 1999-01-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (192)
; US-09-233-506-9
;
; Query Match
; Best Local Similarity 77.1%; Score 16.2; DB 3; Length 192;
; Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
; QY 1 ACCAGATGGAGGACGCAATG 21
; Db 59 ACCACCATGGAGTACGCAATG 39
;
; RESULT 5
; US-09-489-039A-1972/c
; Sequence 1972, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1972
; LENGTH: 1248
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-1972
;
; Query Match
; Best Local Similarity 77.1%; Score 16.2; DB 4; Length 1248;
; Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
; QY 1 ACCAGATGGAGGACGCAATG 21
; Db 437 ACCGGAATGGAGAACCAATG 417
```

## RESULT 6

US-09-949-016-17486  
; Sequence 17486, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17486  
; LENGTH: 56976  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(56976)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17486

Query Match 77.1%; Score 16.2; DB 4; Length 56976;  
Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGAGCAATG 21

Db 13869 ACCAGATGGAGGAGCAATG 13889

## RESULT 7

US-09-949-016-12033  
; Sequence 12033, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12033  
; LENGTH: 101674  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(101674)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12033

Query Match 77.1%; Score 16.2; DB 4; Length 101674;  
Best Local Similarity 85.7%; Pred. No. 3e+02;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGAGCAATG 21

||||| ||||||| |||||

Db 13868 ACCAGTATGGAGGAGCAATG 13888

## RESULT 8

US-08-991-789A-278  
; Sequence 278, Application US/08991789A  
; Patent No. 6225054  
; GENERAL INFORMATION:  
; APPLICANT: Frudakis, Tony N.  
; Reed, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; NUMBER OF SEQUENCES: 292  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group  
; STREET: 701 Fifth Avenue, Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/991,789A  
; FILING DATE: 11-Dec-1997  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Potter, Jane E. R.  
; REGISTRATION NUMBER: 33,332  
; REFERENCE/DOCKET NUMBER: 210121.419C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 278:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 452 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 278:  
US-08-991-789A-278

Query Match 75.2%; Score 15.8; DB 3; Length 452;  
Best Local Similarity 89.5%; Pred. No. 2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCAGATGGAGGAGCAAT 20

Db 109 CCAGATGGAGGAGCAAT 127

## RESULT 9

US-09-062-451-278  
; Sequence 278, Application US/09062451  
; Patent No. 6344550  
; GENERAL INFORMATION:  
; APPLICANT: Frudakis, Tony N.  
; Reed, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; NUMBER OF SEQUENCES: 297  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092



US-09-949-295-278

Query Match 75.2%; Score 15.8; DB 4; Length 452;  
Best Local Similarity 89.5%; Pred. No. 2.e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCAGAATGGAGGCGCAAT 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 109 CCAGAATGGAGGCTGCAAT 127

RESULT 13

US-09-949-016-81705/c  
; Sequence 81705, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 81705  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-81705

Query Match 75.2%; Score 15.8; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 2.e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGCGCAA 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 544 ACCAGAATGGAGGCGCAA 526

RESULT 14

US-09-949-016-81706/c  
; Sequence 81706, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 81706  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-81706

Query Match 75.2%; Score 15.8; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 2.e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGCGCAA 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 169 ACCAGAATGGAGGCGCAA 151

RESULT 15

US-09-949-016-81707/c  
; Sequence 81707, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 81707  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-81707

Query Match 75.2%; Score 15.8; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 2.e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGCGCAA 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 56 ACCAGAATGGAGGCGCAA 38

Search completed: May 16, 2005, 06:11:40  
Job time : 81.5859 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 02:00:36 ; Search time 995.531 Seconds  
(without alignments)  
129.102 Million cell updates/sec

Title: US-10-808-187A-2474

Perfect score: 21

Sequence: 1 accagaatggagcgaatg 21

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/prodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/prodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	21	19	US-10-808-187-2474
2	21	100.0	1620	19	Sequence 2474, Ap
3	21	100.0	2304	19	Sequence 16, Appl
4	21	100.0	2304	19	Sequence 7, Appl
5	21	100.0	24774	19	Sequence 11, Appl
6	21	100.0	28920	19	Sequence 3, Appl
7	21	100.0	28920	19	Sequence 5, Appl
8	21	100.0	29231	19	Sequence 6, Appl
9	21	100.0	29430	19	Sequence 4, Appl
10	21	100.0	29727	18	Sequence 7, Appl
11	21	100.0	29727	18	Sequence 15, Appl
					Sequence 1, Appl

12	21	100.0	29727	19	US-10-889-447-8	Sequence 8, Appl
13	21	100.0	29727	19	US-10-699-936-1	Sequence 1, Appl
14	21	100.0	29736	18	US-10-839-729-17	Sequence 17, Appl
15	21	100.0	29736	19	US-10-889-447-9	Sequence 9, Appl
16	21	100.0	29736	19	US-10-699-936-3	Sequence 3, Appl
17	21	100.0	29742	18	US-10-839-729-16	Sequence 16, Appl
18	21	100.0	29742	19	US-10-808-187-15	Sequence 15, Appl
19	21	100.0	29742	19	US-10-808-187-16	Sequence 16, Appl
20	21	100.0	29742	19	US-10-808-187-240	Sequence 240, App
21	21	100.0	29742	19	US-10-808-187-737	Sequence 737, App
22	21	100.0	29742	19	US-10-808-187-1108	Sequence 1108, Ap
C 23	21	100.0	29742	19	US-10-808-187-1590	Sequence 1590, Ap
C 24	21	100.0	29742	19	US-10-808-187-1965	Sequence 1965, Ap
25	21	100.0	29742	19	US-10-889-447-10	Sequence 10, Appl
C 26	21	100.0	29751	18	US-10-839-729-14	Sequence 14, Appl
27	21	100.0	29751	19	US-10-856-529-1	Sequence 1, Appl
28	21	100.0	29751	19	US-10-626-879-67	Sequence 67, Appl
29	21	100.0	29751	19	US-10-889-447-1	Sequence 1, Appl
30	21	100.0	29751	19	US-10-889-447-2	Sequence 2, Appl
31	21	100.0	29751	19	US-10-699-936-2	Sequence 2, Appl
32	16.8	80.0	701	18	US-10-425-115-37958	Sequence 37958, A
33	16.8	80.0	2499	17	US-10-424-599-119318	Sequence 119318,
34	16.8	80.0	3087	13	US-10-027-632-115311	Sequence 115311,
35	16.8	80.0	3087	13	US-10-027-632-115312	Sequence 115312,
36	16.8	80.0	3087	17	US-10-027-632-115311	Sequence 115311,
37	16.8	80.0	3087	17	US-10-027-632-115312	Sequence 115312,
38	16.8	80.0	3088	13	US-10-027-632-112228	Sequence 112228,
39	16.8	80.0	3088	13	US-10-027-632-112229	Sequence 112229,
40	16.8	80.0	3088	17	US-10-027-632-112228	Sequence 112228,
41	16.8	80.0	3088	17	US-10-027-632-112229	Sequence 112229,
42	16.8	80.0	10419	13	US-10-027-806-3	Sequence 3, Appl
43	16.8	80.0	10419	13	US-10-034-623-3	Sequence 3, Appl
44	16.8	80.0	10419	14	US-10-027-801-3	Sequence 3, Appl
45	16.8	80.0	10419	16	US-10-029-120-3	Sequence 3, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-808-187-2474  
; Sequence 2474, Application US/10808187  
; Publication No. US2005000909A1  
; GENERAL INFORMATION:  
; APPLICANT: PEIRIS, JOSEPH S. M.  
; APPLICANT: YUEN, KWOK YUNG  
; APPLICANT: POON, LIT MAN  
; APPLICANT: GUAN, YI  
; APPLICANT: CHAN, KWOK HUNG  
; APPLICANT: NICHOLLS, JOHN  
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE  
; FILE REFERENCE: V9661.0078  
; CURRENT APPLICATION NUMBER: US/10/808,187  
; CURRENT FILING DATE: 2004-03-24  
; PRIOR APPLICATION NUMBER: 60/457,031  
; PRIOR FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: 60/457,730  
; PRIOR FILING DATE: 2003-03-26  
; PRIOR APPLICATION NUMBER: 60/459,931  
; PRIOR FILING DATE: 2003-04-02  
; PRIOR APPLICATION NUMBER: 60/460,357  
; PRIOR FILING DATE: 2003-04-03  
; PRIOR APPLICATION NUMBER: 60/461,265  
; PRIOR FILING DATE: 2003-04-08  
; PRIOR APPLICATION NUMBER: 60/462,805  
; PRIOR FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 60/468,139  
; PRIOR FILING DATE: 2003-05-05  
; PRIOR APPLICATION NUMBER: 60/464,886  
; PRIOR FILING DATE: 2003-04-23  
; PRIOR APPLICATION NUMBER: 60/471,200  
; PRIOR FILING DATE: 2003-05-16

```
; NUMBER OF SEQ ID NOS: 2476
; SOFTWARE: PatentIn ver. 3.2
; SEQ ID NO 2474
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-808-187-2474
```

```
Query Match      100.0%; Score 21; DB 19; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;
```

```
QY 1 ACCAGAATGGAGCGCAATG 21
    |||||||
DB 1 ACCAGAATGGAGCGCAATG 21
```

```
RESULT 2
US-10-699-936-16
; Sequence 16, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: SARS-CoV ZJ-HZ01
US-10-699-936-16
```

```
Query Match      100.0%; Score 21; DB 19; Length 1620;
Best Local Similarity 100.0%; Pred. No. 2.9; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;
```

```
QY 1 ACCAGAATGGAGCGCAATG 21
    |||||||
DB 157 ACCAGAATGGAGCGCAATG 177
```

```
RESULT 3
US-10-699-936-7
; Sequence 7, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 2304
; TYPE: DNA
; ORGANISM: SARS-CoV Shanghai LY
US-10-699-936-7
```

```
Query Match      100.0%; Score 21; DB 19; Length 2304;
Best Local Similarity 100.0%; Pred. No. 2.9; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;
```

```
QY 1 ACCAGAATGGAGCGCAATG 21
    |||||||
DB 837 ACCAGAATGGAGCGCAATG 857
```

```
RESULT 4
US-10-699-936-11
; Sequence 11, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 2304
; TYPE: DNA
; ORGANISM: SARS coronavirus Shanghai LY
US-10-699-936-11
```

```
Query Match      100.0%; Score 21; DB 19; Length 2304;
Best Local Similarity 100.0%; Pred. No. 2.9; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;
```

```
QY 1 ACCAGAATGGAGCGCAATG 21
    |||||||
DB 837 ACCAGAATGGAGCGCAATG 857
```

```
RESULT 5
US-10-889-447-3
; Sequence 3, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 24774
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-3
```

```
Query Match      100.0%; Score 21; DB 19; Length 24774;
Best Local Similarity 100.0%; Pred. No. 2.7; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;
```

```
QY 1 ACCAGAATGGAGCGCAATG 21
    |||||||
```



```
Db      23315 ACCAGAATGGAGCGCAATG 23335

RESULT 6
US-10-889-447-5
; Sequence 5, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; PRIOR FILING DATE: 2004-07-12
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ03
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-5

Query Match      100.0%; Score 21; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCAGAATGGAGCGCAATG 21
      |||||||
Db      27464 ACCAGAATGGAGCGCAATG 27484

RESULT 7
US-10-889-447-6
; Sequence 6, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; PRIOR FILING DATE: 2004-07-12
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ04
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-6

Query Match      100.0%; Score 21; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCAGAATGGAGCGCAATG 21
      |||||||
Db      27464 ACCAGAATGGAGCGCAATG 27484

RESULT 8
US-10-889-447-4
; Sequence 4, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; PRIOR FILING DATE: 2004-07-12
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 29291
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ02
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-4

Query Match      100.0%; Score 21; DB 19; Length 29291;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCAGAATGGAGCGCAATG 21
      |||||||
Db      27832 ACCAGAATGGAGCGCAATG 27852

RESULT 9
US-10-889-447-7
; Sequence 7, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; PRIOR FILING DATE: 2004-07-12
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 29430
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate GZ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-7

Query Match      100.0%; Score 21; DB 19; Length 29430;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCAGAATGGAGCGCAATG 21
      |||||||
Db      27967 ACCAGAATGGAGCGCAATG 27987

RESULT 10
US-10-839-729-15
; Sequence 15, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
```

; TITLE OF INVENTION: AND METHODS OF USE  
; FILE REFERENCE: BIOBANK.013A  
; CURRENT APPLICATION NUMBER: US/10/839,729  
; CURRENT FILING DATE: 2004-05-04  
; PRIOR APPLICATION NUMBER: 60/468703  
; PRIOR FILING DATE: 2003-05-06  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 29727  
; TYPE: DNA  
; ORGANISM: SARS Coronavirus  
US-10-839-729-15

Query Match 100.0%; Score 21; DB 18; Length 29727;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGACGCAATG 21  
|||||  
DB 28202 ACCAGATGGAGGACGCAATG 28222

RESULT 11  
US-10-827-757-1  
; Sequence 1, Application US/10827757  
; Publication No. US20050004071A1  
; GENERAL INFORMATION:  
; APPLICANT: Comper, Wayne  
; TITLE OF INVENTION: Charged Polysaccharides Resistant To Lysosomal Degradation During  
; TITLE OF INVENTION: Kidney Filtration And Renal Passage And Their Use To Treat Or  
; TITLE OF INVENTION: Prevent Infection By Coronaviruses

; FILE REFERENCE: 11213-007-999  
; CURRENT APPLICATION NUMBER: US/10/827,757  
; CURRENT FILING DATE: 2004-04-20  
; PRIOR APPLICATION NUMBER: 60/464,294  
; PRIOR FILING DATE: 2003-04-21  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 29727  
; TYPE: DNA  
; ORGANISM: SARS-related coronavirus (Urbani strain)  
US-10-827-757-1

Query Match 100.0%; Score 21; DB 18; Length 29727;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGACGCAATG 21  
|||||  
DB 28202 ACCAGATGGAGGACGCAATG 28222

RESULT 12  
US-10-889-447-8  
; Sequence 8, Application US/10889447  
; Publication No. US20050075307A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Jain, Ravi  
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
; FILE REFERENCE: RTS-0685US  
; CURRENT APPLICATION NUMBER: US/10/889,447  
; CURRENT FILING DATE: 2004-07-12  
; PRIOR APPLICATION NUMBER: 60/486,670  
; PRIOR FILING DATE: 2003-07-12  
; NUMBER OF SEQ ID NOS: 241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 29727  
; TYPE: DNA  
; ORGANISM: SARS coronavirus Urbani

US-10-889-447-8

Query Match 100.0%; Score 21; DB 19; Length 29727;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGACGCAATG 21  
|||||  
DB 28202 ACCAGATGGAGGACGCAATG 28222

RESULT 13  
US-10-699-936-1  
; Sequence 1, Application US/10699936  
; Publication No. US20050095582A1  
; GENERAL INFORMATION:  
; APPLICANT: Gillim-Ross, Laura  
; APPLICANT: Taylor, Jill  
; APPLICANT: Scholl, David R.  
; APPLICANT: Wentworth, David E.  
; APPLICANT: Jollick, Joseph D.  
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory  
; TITLE OF INVENTION: Syndrome Coronavirus  
; FILE REFERENCE: DHI-07986  
; CURRENT APPLICATION NUMBER: US/10/699,936  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 29727  
; TYPE: DNA  
; ORGANISM: SARS coronavirus Urbani  
US-10-699-936-1

Query Match 100.0%; Score 21; DB 19; Length 29727;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGACGCAATG 21  
|||||  
DB 28202 ACCAGATGGAGGACGCAATG 28222

RESULT 14  
US-10-839-729-17  
; Sequence 17, Application US/10839729  
; Publication No. US20050002953A1  
; GENERAL INFORMATION:  
; APPLICANT: Jens Herold  
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES  
; TITLE OF INVENTION: AND METHODS OF USE  
; FILE REFERENCE: BIOBANK.013A  
; CURRENT APPLICATION NUMBER: US/10/839,729  
; CURRENT FILING DATE: 2004-05-04  
; PRIOR APPLICATION NUMBER: 60/468703  
; PRIOR FILING DATE: 2003-05-06  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 29736  
; TYPE: DNA  
; ORGANISM: SARS Coronavirus  
US-10-839-729-17

Query Match 100.0%; Score 21; DB 18; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGACGCAATG 21  
|||||  
DB 28187 ACCAGATGGAGGACGCAATG 28207

RESULT 15  
US-10-889-447-9  
; Sequence 9, Application US/10889447  
; Publication No. US20050075307A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Jain, Ravi  
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
; FILE REFERENCE: RTS-068505  
; CURRENT APPLICATION NUMBER: US/10/889,447  
; CURRENT FILING DATE: 2004-07-12  
; PRIOR APPLICATION NUMBER: 60/486,670  
; PRIOR FILING DATE: 2003-07-12  
; NUMBER OF SEQ ID NOS: 241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 29736  
; TYPE: DNA  
; ORGANISM: SARS coronavirus CUHK-W1  
US-10-889-447-9

Query Match 100.0%; Score 21; DB 19; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCAGAATGGAGCGCAATG 21  
|||||  
Db 28187 ACCAGAATGGAGCGCAATG 28207

Search completed: May 16, 2005, 14:33:49  
Job time : 997.531 secs

**THIS PAGE BLANK (USPTO)**



by N. Garrett, K. Ryan and A.M. Zorn, (Wellcome/CRC Institute). "

## ORIGIN

Query Match 92.4%; Score 19.4; DB 5; Length 470;  
Best Local Similarity 95.2%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGACCAATG 21

|||||

Db 418 ACCAGAATGGAGGACCAATG 398

## RESULT 2

BJ058057/c  
LOCUS  
DEFINITION BJ058057 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL058c18 5', mRNA sequence.

ACCESSION BJ058057

VERSION BJ058057.1 GI:17470401

KEYWORDS

SOURCE xenopus laevis (African clawed frog)

ORGANISM

xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
Xenopodinae; Xenopus; Xenopus.

## REFERENCE

AUTHORS Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.

TITLE Expressed genes in X. laevis embryo

JOURNAL Unpublished (2001)

COMMENT Contact: Tadabu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tsini@genes.nig.ac.jp

The information of this clone is available through the following URL.

http://xenopus.nibb.ac.jp.

## FEATURES

source

1..566  
Location/Qualifiers  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="XL058c18"  
/tissue\_type="whole embryo"  
/dev\_stage="stage 25"  
/clone\_lib="NIBB Mochii normalized Xenopus tailbud library"

## ORIGIN

Query Match 92.4%; Score 19.4; DB 4; Length 566;  
Best Local Similarity 95.2%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGACCAATG 21

|||||

Db 437 ACCAGAATGGAGGACCAATG 417

## RESULT 3

BG160575/c  
LOCUS  
DEFINITION BG160575 Wellcome CRC prn3 head Xenopus laevis cDNA clone IMAGE:3580537 5', mRNA sequence.

ACCESSION BG160575

VERSION BG160575.1 GI:12694494

KEYWORDS

SOURCE xenopus laevis (African clawed frog)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## REFERENCE

AUTHORS

## TITLE

JOURNAL

COMMENT

Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
Xenopodinae; Xenopus; Xenopus.

1 (bases 1 to 667)

Clifton, S., Johnson, S.D., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.

WashU Xenopus EST project, 1999

Unpublished (1999)

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by N. Garrett, K. Ryan and A.M. Zorn,

(Wellcome/CRC Institute). DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov

Seq primer: -40RP from Gibco

High quality sequence stop: 508.

## FEATURES

source

1..667  
Location/Qualifiers  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="IMAGE:3580537"  
/tissue\_type="head, stage 30"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Wellcome CRC prn3 head"  
/note="Vector: pBSRN3; Site\_1: NotI; Site\_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library was constructed by N. Garrett, K. Ryan and A.M. Zorn, (Wellcome/CRC Institute)."

## ORIGIN

Query Match 92.4%; Score 19.4; DB 4; Length 667;  
Best Local Similarity 95.2%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGACCAATG 21

|||||

Db 493 ACCAGAATGGAGGACCAATG 473

## RESULT 4

FI5026/c  
LOCUS  
DEFINITION F15026 SSO4H12 Porcine small intestine cDNA library Sus scrofa cDNA clone O4h12, mRNA sequence.

ACCESSION F15026

VERSION F15026.1 GI:972497

KEYWORDS

SOURCE Sus scrofa (pig)

ORGANISM

Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 286)

Wintero, A.K., Fredholm, M. and Davies, W.

Evaluation and characterization of a porcine small intestine cDNA

library: analysis of 839 clones

Mamm. Genome 7 (7), 509-517 (1996)

96327607

8672129

Contact: A.K. Wintero

Department of Animal Science and Animal Health, Division of Animal Genetics, The Royal Veterinary and Agricultural University

Bulowsvej 13, 1870 Frederiksberg C, Denmark.

FEATURES

source

1..286

```

/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:8823"
/clone="o4h12"
/clone_lib="Porcine small intestine cDNA library"
/note="directionally cloned cDNA in XLI-blue MRF"

ORIGIN
Query Match      84.8%; Score 17.8; DB 7; Length 286;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGACCGCAATG 21
|||||
Db 272 ACCAGAATGGGGATGCAATG 252

RESULT 5
BG893390/c
LOCUS
DEFINITION
dae20b11.x1 NICHDD_XGC_Lul Xenopus laevis cDNA clone IMAGE:4674405
3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 367)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
High quality sequence stop: 344.
Location/Qualifiers
1. 367
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:4674405"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDD_XGC_Lul"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.6 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection
(XGC) library."

ORIGIN
Query Match      84.8%; Score 17.8; DB 4; Length 367;
Best Local Similarity 90.5%; Pred. No. 6.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGATGGAGGACCGCAATG 21
|||||
Db 358 ACCAGAATGGGGACCGCAATG 338

RESULT 6
BP677561
LOCUS
DEFINITION
BP677561 Osada Taira anterior neuroectoderm (ANE) pCS105 cDNA

/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:8823"
/clone="o4h12"
/clone_lib="Porcine small intestine cDNA library"
/note="directionally cloned cDNA in XLI-blue MRF"

```

```

library Xenopus laevis cDNA clone XL427d08ex 5', mRNA sequence.
BP677561
LOCUS
DEFINITION
BP677561.1 GI:46025516
EST.
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 406)
Osada,S., Kitayama,A., Ueno,N. and Taira,M.
Expression analysis of genes which are expressed in the anterior
neuroectoderm of Xenopus embryos
Unpublished (2004)
Contact: Masanori Taira
Department of Biological Sciences
Graduate School of Science, University of Tokyo; CREST, Japan
Science and Technology Corporation, Japan
7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-0033, Japan
Tel: 81-03-5841-4434
Fax: 81-03-5841-4434
Email: m.taira@biol.s.u-tokyo.ac.jp,
URL:http://www.shigen.nig.ac.jp/nbrp/xenopus/est/.
Location/Qualifiers
1. 406
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL427d08ex"
/tissue_type="anterior neuroectoderm"
/dev_stage="late gastrula (stage 12.5)"
/clone_lib="Osada Taira anterior neuroectoderm (ANE)
pCS105 cDNA library"

ORIGIN
Query Match      84.8%; Score 17.8; DB 5; Length 406;
Best Local Similarity 90.5%; Pred. No. 6.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCAGAATGGAGGACCGCAATG 21
|||||
Db 189 ACCAGAATGGGGACCGCAATG 209

RESULT 7
CF916257
LOCUS
DEFINITION
CF916257-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus
musculus cDNA clone NIA:B0992E02 IMAGE:30481105 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 407)
Piao,Y., KO,N.T., Lim,M.K. and KO,M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
JOURNAL
MEDLINE
PUBMED
21429098
11544199
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: B0992 row: E column: 02
Seq primer: M13 Reverse
High quality sequence stop: 407
POLYA=No.

```

## FEATURES

## source

## Location/Qualifiers

```

1..407
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:B0992E02-5"
/db_xref="taxon:10090"
/clone="NIA:B0992E02 IMAGE:30481105"
/dev_stage="Unfertilized Egg"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long
1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were extracted from a pool of 1488 unfertilized eggs.
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer [Invitrogen]:
5'-pGACTAGTTCCTAGATCGGAGCGGCCCTTTTTTTTTTTT-3',
treated with T4 DNA polymerase, and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker LL-SalI, purified by phenol/chloroform, and
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer SalI-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.5 kb. The library was
constructed by Yulan Piao."

```

## ORIGIN

```

Query Match      84.8%; Score 17.8; DB 7; Length 407;
Best Local Similarity 90.5%; Pred. No. 6.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ACCAGATGGAGGAGCGCAATG 21

DB 385 AACAAATGGAGGAGCGCAATG 405

## RESULT 8

## CO296590/c

## LOCUS

```

DEFINITION      433 bp mRNA linear EST 25-JUN-2004
melanogaster cDNA clone EK214255 5, mRNA sequence.

```

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

```

Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

```

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

```

Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: EK.2142 row: E column: 7
High quality sequence stop: 432.
Location/Qualifiers

```

## FEATURES

## source

```

1..433
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="EK214255"
/clone_lib="Exelixis FlyTag CK01 pCDNA-SK+"
/note="Organ: mixed stage embryos, imaginal disks, and
adult heads; Vector: pCDNA-SK+; Site_1: NotI; Site_2:
XhoI; Random primed, normalized library from mixed stage
embryos, imaginal disks, and adult heads."

```

## ORIGIN

```

Query Match      84.8%; Score 17.8; DB 7; Length 433;
Best Local Similarity 90.5%; Pred. No. 6.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ACCAGATGGAGGAGCGCAATG 21

DB 58 ACCAGATGGAGGAGCGCAATG 38

## RESULT 9

## BI446008

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis

## Xenopus laevis



**BG635276**  
**LOCUS**  
**DEFINITION** BG635276 577 bp mRNA linear EST 23-APR-2001  
 AT31804.5:prime AT Drosophila melanogaster adult testes pOTB7  
 Drosophila melanogaster cDNA clone AT31804 5 similar to CG3845:  
 FBan0003845 located on: 2R 49E1-49E1.; 04/10/2001, mRNA sequence.  
**ACCESSION**  
**VERSION** BG635276.1 GI:13762813  
**KEYWORDS**  
**SOURCE** EST.  
**ORGANISM** Drosophila melanogaster (fruit fly)  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
**REFERENCE**  
**AUTHORS** 1 (bases 1 to 577)  
 Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,  
 Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,  
 Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,  
 Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J.,  
 Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,  
 Lewis, S.B., Celnikier, S. and Rubin, G.M.  
**TITLE** BDGP/HMI AT Drosophila EST Project  
**JOURNAL** Unpublished (2000)  
**COMMENT** Contact: Stapleton, M.  
 BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
 hit genomic AE003820: arm:2R [7816709,8082095]  
 estimated-cyto:49D1-49F2: 04/10/2001  
 Plate: AT.318 row: A column: 4  
 High quality sequence stop: 574.  
**FEATURES**  
 source  
 1..577  
 Location/Qualifiers  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"  
 /clone="AT31804"  
 /sex="male"  
 /dev\_stage="0-3 day old Ore-R males"  
 /lab\_host="Plates AT.10-AT.120: DH5-alpha. Plates  
 AT.121-AT.319: DH5-alpha Tona"  
 /clone\_lib="AT Drosophila melanogaster adult testes pOTB7"  
 /note="Organ: ADULT testes; Vector: pOTB7; Site 1: EcoRI;  
 Site 2: XhoI; The mRNA for the testis library was made  
 from testes and seminal vesicles hand dissected from 0-3  
 day old Ore-R males. RNA kindly provided by the lab of  
 Margaret Fuller. Sized fractionated cDNAs were directly  
 ligated into pOTB7. Plasmid cDNA library."  
**ORIGIN**  
 Query Match 84.8%; Score 17.8; DB 4; Length 577;  
 Best Local Similarity 90.5%; Pred. No. 6.7e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
**QY** 1 ACCAGAATGGAGGAGCGCAATG 21  
 |||||  
**Db** 105 ACCAGAATGGAGGAGCGCAATG 125  
 |||||  
**RESULT 11**  
**AW640026**  
**LOCUS** AW640026 585 bp mRNA linear EST 26-APR-2001  
**DEFINITION** bl89h06 w1 Blackshear/Soares normalized Xenopus egg library Xenopus  
 laevis cDNA clone PBX0089H06 5', mRNA sequence.  
**ACCESSION**  
**VERSION** AW640026.1 GI:7397236  
**KEYWORDS** EST.  
**SOURCE** Xenopus laevis (African clawed frog)  
 Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus; Xenopus.

**REFERENCE**  
**AUTHORS** 1 (bases 1 to 585)  
 Blackshear, P.J., Lai, W.S., Thorn, J.M., Kennington, E.A., Staffa, N.G.,  
 Jr., Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M.,  
 Touchman, J.W., Bonaldo, M.P. and Soares, M.B.  
 The NIEHS Xenopus maternal EST project: interim analysis of the  
 first 13,879 ESTs from unfertilized eggs  
 Gene 267 (1), 71-87 (2001)  
**JOURNAL** 21211403  
**MEDLINE** 11311557  
**PUBMED**  
**COMMENT** Contact: Perry J. Blackshear  
 Office of Clinical Research and Laboratory of Signal Transduction  
 National Institute of Environmental Health Sciences  
 A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,  
 USA  
 Tel: 919 541-4899  
 Fax: 919 541-4571  
 Email: [black000@niehs.nih.gov](mailto:black000@niehs.nih.gov)  
 Clone is available through Research Genetics, Inc., 2130 Memorial  
 Parkway, Huntsville, AL 35901  
 phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email  
 cdna@resgen.com  
 DNA Sequencing and analyses performed by National Institutes of  
 Health Intramural Sequencing Center (NISC).  
 PCR Primers  
 FORWARD: TGTAAACGACGCGCCAGT  
 BACKWARD: CAGGAACAGCTATGACC  
 Plate: 0089 row: H column: 06  
 Seq primer: T7 primer.  
**FEATURES**  
 source  
 1..585  
 Location/Qualifiers  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="PBX0089H06"  
 /sex="female"  
 /tissue\_type="unfertilized egg"  
 /cell\_type="unfertilized egg"  
 /dev\_stage="unfertilized egg"  
 /lab\_host="DH10B"  
 /clone\_lib="Blackshear/Soares normalized Xenopus egg  
 library"  
 /note="Vector: pT7T3-Pac; Site 1: EcoRI; Site 2: NotI;  
 PolyA-selected mRNA was prepared from unfertilized Xenopus  
 laevis eggs. The library was constructed in the vector  
 pT7T3-Pac as described in Bonaldo, M.P., Lennon, G. and  
 Soares, M.B. 'Normalization and subtraction: two  
 approaches to facilitate gene discovery', Genome Research  
 6:791-806, 1996. The first strand synthesis used a  
 NotI-dri18 primer; double stranded cDNAs were ligated to  
 EcoRI adapters, digested with NotI, and directionally  
 cloned into the NotI and EcoRI-digested pT7T3-Pac vector.  
 The library contained approximately 7.2 x 10<sup>5</sup>  
 recombinants, with average insert sizes of 1-1.5 kb."  
**ORIGIN**  
 Query Match 84.8%; Score 17.8; DB 2; Length 585;  
 Best Local Similarity 90.5%; Pred. No. 6.7e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
**QY** 1 ACCAGAATGGAGGAGCGCAATG 21  
 |||||  
**Db** 318 ACCAGAATGGAGGAGCGCAATG 338  
 |||||  
**RESULT 12**  
**CF610686/c**  
**LOCUS** CF610686 597 bp mRNA linear EST 30-SEP-2003  
**DEFINITION** Lr Cd2CF 13G04 M13R Earthworm Cadmium Exposure Library Lumbricus  
 rubellus cDNA clone Lr\_Cd2CF\_13G04 5', mRNA sequence.  
**ACCESSION**  
**VERSION** CF610686.1 GI:37191864  
**KEYWORDS** EST.  
**SOURCE** Lumbricus rubellus (humus earthworm)



RESULT 15  
BJ078249/c  
LOCUS  
DEFINITION BJ078249 NIBB Mochii normalized Xenopus tailbud library Xenopus  
laevis cDNA clone XL063j06 3', mRNA sequence.  
ACCESSION  
VERSION BJ078249.1 GI:17523165  
KEYWORDS EST.  
SOURCE  
ORGANISM Xenopus laevis (African clawed frog)  
Xenopus laevis  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
Xenopodinae; Xenopus; Xenopus.  
REFERENCE  
1 (bases 1 to 607)  
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and  
Kohara,Y.  
TITLE Expressed genes in X. laevis embryo  
JOURNAL Unpublished (2001)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp  
The information of this clone is available through the following  
URL.  
http://xenopus.nibb.ac.jp.  
FEATURES  
source  
1..607  
Location/Qualifiers  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="XL063j06"  
/tissue\_type="whole embryo"  
/dev\_stage="stage 25"  
/clone\_lib="NIBB Mochii normalized Xenopus tailbud  
library"  
ORIGIN  
Query Match 84.8%; Score 17.8; DB 4; Length 607;  
Best Local Similarity 90.5%; Pred. No. 6.8e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ACCAGATGGAGGACGCAATG 21  
|||||  
Db 170 ACCAGATGGGGGACCCCAATG 150  
|||||

Search completed: May 16, 2005, 06:03:36  
Job time : 2379.69 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 21:16:45 ; Search time 597.656 Seconds  
(without alignments)  
2026.885 Million cell updates/sec

Title: US-10-808-187A-2475

Perfect score: 25  
Sequence: 1 gctgtgaaccaagacgcagattatt 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25	100.0	1269	AY360146	AY360146 SARS coro
C 2	25	100.0	1269	AY365036	AY365036 SARS coro
C 3	25	100.0	1669	AY536760	AY536760 SARS coro
C 4	25	100.0	2304	AY322205S4	AY322208 SARS coro
C 5	25	100.0	2810	AY290752	AY290752 SARS coro
C 6	25	100.0	3080	AY443086S10	AY443095 SARS coro
C 7	25	100.0	8439	AY304489	AY304489 SARS coro
C 8	25	100.0	8581	AY304487	AY304487 SARS coro
C 9	25	100.0	11006	AY304491	AY304491 SARS coro
C 10	25	100.0	11010	AY304493	AY304493 SARS coro
C 11	25	100.0	11010	AY304494	AY304494 SARS coro
C 12	25	100.0	13471	AY304490	AY304490 SARS coro
C 13	25	100.0	13471	AY304492	AY304492 SARS coro
C 14	25	100.0	29013	AY463060	AY463060 SARS coro
C 15	25	100.0	29350	AY394999	AY394999 SARS coro
C 16	25	100.0	29350	AY395000	AY395000 SARS coro
C 17	25	100.0	29350	AY395001	AY395001 SARS coro
C 18	25	100.0	29350	AY395002	AY395002 SARS coro
C 19	25	100.0	29433	AY394977	AY394977 SARS coro

C 20	25	100.0	29530	14	AY394985	AY394985 SARS coro
C 21	25	100.0	29573	14	AY338174	AY338174 SARS coro
C 22	25	100.0	29573	14	AY338175	AY338175 SARS coro
C 23	25	100.0	29573	14	AY348314	AY348314 SARS coro
C 24	25	100.0	29577	14	AY559094	AY559094 SARS coro
C 25	25	100.0	29592	14	AY463059	AY463059 SARS coro
C 26	25	100.0	29620	14	AY395004	AY395004 SARS coro
C 27	25	100.0	29640	14	AY394978	AY394978 SARS coro
C 28	25	100.0	29645	14	AY394979	AY394979 SARS coro
C 29	25	100.0	29646	14	AY394982	AY394982 SARS coro
C 30	25	100.0	29647	14	AY395003	AY395003 SARS coro
C 31	25	100.0	29661	14	AY559086	AY559086 SARS coro
C 32	25	100.0	29685	14	AY394988	AY394988 SARS coro
C 33	25	100.0	29670	14	AY559082	AY559082 SARS coro
C 34	25	100.0	29683	14	AY394996	AY394996 SARS coro
C 35	25	100.0	29683	14	AY394997	AY394997 SARS coro
C 36	25	100.0	29699	14	AY394983	AY394983 SARS coro
C 37	25	100.0	29705	14	AY283795	AY283795 SARS coro
C 38	25	100.0	29705	14	AY394980	AY394980 SARS coro
C 39	25	100.0	29706	14	AY283797	AY283797 SARS coro
C 40	25	100.0	29709	14	AY394987	AY394987 SARS coro
C 41	25	100.0	29711	14	AY283794	AY283794 SARS coro
C 42	25	100.0	29711	14	AY283796	AY283796 SARS coro
C 43	25	100.0	29711	14	AY427439	AY427439 SARS coro
C 44	25	100.0	29712	14	AY559093	AY559093 SARS coro
C 45	25	100.0	29713	14	AY559085	AY559085 SARS coro

ALIGNMENTS

RESULT 1  
AY360146/c  
LOCUS  
DEFINITION  
SARS coronavirus HPZ-2003 nucleocapsid protein (NP) gene, complete cds.  
ACCESSION  
AY360146  
VERSION  
AY360146.1  
KEYWORDS  
GI:34329618  
SOURCE  
SARS coronavirus HPZ-2003  
ORGANISM  
SARS coronavirus HPZ-2003  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus.  
REFERENCE  
1 (bases 1 to 1269)  
Zhu,H.P., Lu,Q.Y., Yao,P.P., Xu,F., Ge,Q., Weng,J.Q.,  
Yan,J.Y., Gong,L.M., Shi,W., Zhao,Z.Y. and Zhu,Z.Y.  
TITLE  
Cloning and expression of nucleocapsid protein gene of SARS associated coronavirus  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 1269)  
Zhu,H.P., Lu,Q.Y., Yao,P.P., Xu,F., Ge,Q., Weng,J.Q.,  
Yan,J.Y., Gong,L.M., Shi,W., Zhao,Z.Y. and Zhu,Z.Y.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (04-AUG-2003) Zhejiang Provincial Key Laboratory of HFRS, Zhejiang Center for Disease Prevention and Control, 17 Lao Zhe Da Zhi Road, Hangzhou, Zhejiang 310009, P.R. China

FEATURES  
source  
1..1269  
/organism="SARS coronavirus HPZ-2003"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:243925"  
gene  
1..1269  
/gene="NP"  
CDS  
1..1269  
/gene="NP"  
/codon\_start=1  
/product="nucleocapsid protein"  
/protein\_id="AA063890.1"  
/db\_xref="GI:34329619"  
/translation="MSDNGPQSNQSRAPRITFCGPTDSTDNQNGNGARPQRPPQ  
GLPNNTASFTALTQHGKEELFRGQGVPIINTSGDDQIGYIRATRRVRGDDGQM  
KELSPWYFYLTGPEASLPGANKEGIWVATEGALNTPKDHIQTRNPNNAATVL  
QLPQGTTLPGFVAEGSRGGSQSSRSSRSRNRNSTPGSSRGNSPARMASGGGET

ALALLLLRLNQLSKSVKGGQOQOQGVTTKKSAEASKKPKRKTATKQYNVTOAFG  
RRGPQTQNGFGDQLIRQGYDKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY  
HGAIKLDDKDPQKDNVILLNKHIDAYKTFFPTEPKDKKKKTDEAQLPQQRKKQPT  
VTLLPAADMDDFSRLQNSMSGASADSTQA"

ORIGIN  
Query Match 100.0%; Score 25; DB 14; Length 1269;  
Best Local Similarity 100.0%; Pred. No. 0.033; 0; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

QY 1 GCTGTGAACCAAGACGACGAGTATTAT 25  
|||||  
Db 167 GCTGTGAACCAAGACGACGAGTATTAT 143

RESULT 2  
LOCUS AY365036/c 1269 bp RNA linear VRL 01-JUN-2004  
DEFINITION SARS coronavirus HB nucleocapsid protein (NP) gene, complete cds.  
ACCESSION AY365036  
VERSION AY365036.1 GI:38176101  
KEYWORDS  
SOURCE  
ORGANISM

SARS coronavirus HB  
SARS coronavirus HB  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.  
1 (bases 1 to 1269)

REFERENCE  
AUTHORS Timani, K.A., Ye, L., Ye, L., Zhu, Y., Wu, Z. and Gong, Z.  
TITLE Cloning, sequencing, expression, and purification of  
SARS-associated coronavirus nucleocapsid protein for serodiagnosis  
of SARS  
J. Clin. Virol. 30 (4), 309-312 (2004)

J. Clin. Virol. 30 (4), 309-312 (2004)  
PUBMED 15163419  
REFERENCE 2 (bases 1 to 1269)  
AUTHORS Timani, K.A., Ye, L., Ye, L., Zhu, Y. and Guo, D.  
TITLE Direct Submission  
Submitted (11-AUG-2003) Institute of Virology, Wuhan University,  
Wuhan, Hubei 430072, P.R. China

FEATURES  
Location/Qualifiers  
1..1269

/organism="SARS coronavirus HB"  
/viroion

/mol\_type="genomic RNA"

/isolate="HB"

/db\_xref="taxon:253634"

/country="China: Wuhan, Hubei"

1..1269

/gene="NP"

1..1269

/gene="NP"

/note="N protein"

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAR1290.1"

/db\_xref="GI:38176102"

/translation="MSDNGPQNSAPRITFGGPTDSTNNQNGRNARPKRRPQ  
GLPNTASFTALTQHGKEELFRPGQVPINTNSGPDQDQIGYRATRRVRGGDKM  
KELSPRWYFYLTGTPEASLPYCANKEGIWVATEGALNTPKDHIGTRPNNNAATVL  
OLPQGTTLPGFVAGSGSGSOASSRSSRSRSGNSNSTPGSSRSGNSPARMASGGGT  
ALALLLLRLNQLSKSVKGGQOQOQGVTTKKSAEASKKPKRKTATKQYNVTOAFG  
RRGPQTQNGFGDQLIRQGYDKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY  
HGAIKLDDKDPQKDNVILLNKHIDAYKTFFPTEPKDKKKKTDEAQLPQQRKKQPT  
VTLLPAADMDDFSRLQNSMSGASADSTQA"

ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 1269;  
Best Local Similarity 100.0%; Pred. No. 0.033; 0; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

QY 1 GCTGTGAACCAAGACGACGAGTATTAT 25

|||||

Db 167 GCTGTGAACCAAGACGACGAGTATTAT 143

RESULT 3  
LOCUS AY336760/c

DEFINITION SARS coronavirus BJO1 nucleocapsid protein mRNA, complete cds.

ACCESSION AY336760

VERSION AY336760.3 GI:52100973

KEYWORDS

SOURCE

ORGANISM

SARS coronavirus BJO1

SARS coronavirus BJO1

Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

1 (bases 1 to 1669)

Li, T., Li, X., Chang, Z. and Liu, L.

Identification of SARS-CoV mRNA leader sequence

Unpublished

2 (bases 1 to 1669)

Li, T., Li, X., Liu, L. and Chang, Z.

Direct Submission

Submitted (30-JAN-2004) Institute of Biomedicine, Tsinghua

University, Beijing, Beijing 100084, China

3 (bases 1 to 1669)

Li, T., Li, X., Liu, L. and Chang, Z.

Direct Submission

Submitted (06-JUL-2004) Institute of Biomedicine, Tsinghua

University, Beijing, Beijing 100084, China

Sequence update by submitter

4 (bases 1 to 1669)

Li, T., Li, X., Chang, Z. and Liu, L.

Direct Submission

Submitted (15-SEP-2004) Institute of Biomedicine, Tsinghua

University, Beijing 100084, China

Sequence update by submitter

COMMENT On Sep 15, 2004 this sequence version replaced gi:49921010.

FEATURES

Location/Qualifiers

1..1669

/organism="SARS coronavirus BJO1"

/mol\_type="mRNA"

/strain="BJO1"

/db\_xref="taxon:228407"

/country="China"

81..1349

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAS48456.1"

/db\_xref="GI:44889746"

/translation="MSDNGPQNSAPRITFGGPTDSTNNQNGRNARPKRRPQ  
GLPNTASFTALTQHGKEELFRPGQVPINTNSGPDQDQIGYRATRRVRGGDKM  
KELSPRWYFYLTGTPEASLPYCANKEGIWVATEGALNTPKDHIGTRPNNNAATVL  
OLPQGTTLPGFVAGSGSGSOASSRSSRSRSGNSNSTPGSSRSGNSPARMASGGGT  
ALALLLLRLNQLSKSVKGGQOQOQGVTTKKSAEASKKPKRKTATKQYNVTOAFG  
RRGPQTQNGFGDQLIRQGYDKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY  
HGAIKLDDKDPQKDNVILLNKHIDAYKTFFPTEPKDKKKKTDEAQLPQQRKKQPT  
VTPLPAADMDDFSRLQNSMSGASADSTQA"

ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 1669;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGACGAGTATTAT 25

|||||

Db 247 GCTGTGAACCAAGACGACGAGTATTAT 223

RESULT 4

LOCUS AY32220584/c

DEFINITION SARS coronavirus Shanghai Y. Orf7a gene, partial cds; and Orf7b,  
Orf8A, Orf8b, and nucleocapsid protein genes, complete cds.

ACCESSION AY322208

VERSION AY322208.1 GI:32454342

KEYWORDS

```
4 of 4
SEGMENT SARS coronavirus Shanghai LY
SOURCE SARS coronavirus Shanghai LY
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 2304)
AUTHORS Yuan,Z., Zhang,X., Hu,Y., Lan,S., Wang,H., Zhou,Z. and Wen,Y.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2003) Molecular Virology, Shanghai Medical
College of Fudan University, 138 Yi Xue Yuan Road, Shanghai 200032,
P.R. China
FEATURES
source Location/Qualifiers
1..2304
/organism="SARS coronavirus Shanghai LY"
/mol_type="genomic RNA"
/db_xref="taxon:235173"
/country="China: Shanghai"
<1..276
/codon_start=1
/product="Orf7a"
/protein_id="AAP82970.1"
/db_xref="GI:32454348"
/translation="KEPCPSATYEGNSPPHPLADNKFALTCTCTHFAFACADGTRHTY
OLRARSVSPKLPFIRQEEVQQLYPLFLIVALVFLICFTIKRTE"
273..407
/codon_start=1
/product="Orf7b"
/protein_id="AAP82971.1"
/db_xref="GI:32454349"
/translation="MNELTLDIPLYCLFLAPLLFLVLMILIFWFSLEIQDLBEPCTKV"
414..533
/codon_start=1
/product="Orf8a"
/protein_id="AAP82972.1"
/db_xref="GI:32454350"
/translation="MKLLIVLTICISLCICTVQRCASNKPVLBPCVKQH"
499..753
/codon_start=1
/product="Orf8b"
/protein_id="AAP82973.1"
/db_xref="GI:32454351"
/translation="MKLILVRVNTGNITYSTAWLCALGVLPFHRWHVWVQTCTPNV
TINCQDPAGGALIARCWYLHQBQTAARFVVLVLRKTN"
755..2023
/codon_start=1
/product="nucleocapsid protein"
/protein_id="AAP82974.1"
/db_xref="GI:32454352"
/translation="MSDNGPQNSAPRITFGPTDSTDNNGNGRNGARPKQRPPQ
GLPNTASTFTALTQHGKELRPPRGQVPINTSGPDQIQYRRATRRVRGGDKM
KELSPRWYFLGTGPEALPYGANKGIVWVATEGALNTPKHIGTRPNNAATVL
OLPQGTTLPGFVAGSGSGSOASSRSRSGNSNSNTPGSRGNSPARMAGGGGT
ALALLDLRLNQLSKVSGKQQQQCTVTKGSAASAKPKRKTATKQVNTQAFG
RKGPEQTQNGQDQDLIRGTDYKHPWQIAQFAPSASAFGMSRIQMEAPSGTWLTY
HGAIKLDDKDPQKDVLLNKHIDAYKTFPTPEPKDKKKKTDEAQPLPFRQKKQPT
VTLLPAADMDDFSRQLQNSGASADSTQA"
ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 2304;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTGTGACCAACGACGATTATT 25
Db 921 GCTGTGACCAACGACGATTATT 897
RESULT 5
AY290752/c 2810 bp RNA linear VRL 09-FEB-2004
LOCUS SARS coronavirus ZJ01 isolate ZJ01b uncharacterized protein 6,
DEFINITION uncharacterized protein 7a, uncharacterized protein 7b,
```

```
uncharacterized protein 8a, uncharacterized protein 8b,
nucleocapsid protein, uncharacterized protein 9b, and
uncharacterized protein 9c genes, complete cds.
AY290752 GI:38735509
SARS coronavirus ZJ01
SARS coronavirus ZJ01
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 2810)
AUTHORS Li,L., Wang,Z., Lu,Y., Bao,Q., Chen,S., Wu,N., Cheng,S., Weng,J.,
Zhang,Y., Yan,J., Mei,L., Wang,X., Zhu,H., Yu,Y., Zhang,M., Li,M.,
Yao,J., Lu,Q., Yao,P., Bo,X., Wo,J., Wang,S. and Hu,S.
TITLE Severe acute respiratory syndrome-associated coronavirus genotype
and its characterization
JOURNAL Chin. Med. J. 116 (9), 1288-1292 (2003)
MEDLINE 22889812
PUBMED 14527350
REFERENCE 2 (bases 1 to 2810)
Wang,Z.G., Li,L.J., Luo,Y., Zhang,J.Y., Wang,M.Y., Cheng,S.Y.,
Zhang,Y.J., Wang,X.M., Lu,Y.Y., Wu,N.P., Mei,L.L. and Wang,Z.X.
TITLE Molecular biological analysis of genotyping and phylogeny of severe
acute respiratory syndrome associated coronavirus
JOURNAL Chin. Med. J. 117 (1), 42-48 (2004)
PUBMED 14733771
REFERENCE 3 (bases 1 to 2810)
Wang,Z., Cheng,S., Zhang,Y., Lu,Y. and Mei,L.
AUTHORS Direct Submission
TITLE Submitted (03-MAY-2003) Department of Microbiology, Zhejiang Center
of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou,
Zhejiang 310009, China
4 (bases 1 to 2810)
Wang,Z., Cheng,S., Zhang,Y., Lu,Y. and Mei,L.
AUTHORS Direct Submission
TITLE Submitted (05-DEC-2003) Department of Microbiology, Zhejiang Center
of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou,
Zhejiang 310009, China
REMARK Nucleotide and amino acid sequences updated by submitter
COMMENT On Dec 5, 2003 this sequence version replaced gi:31505969.
FEATURES
source
1..2810
/organism="SARS coronavirus ZJ01"
/mol_type="genomic RNA"
/isolate="ZJ01b"
/isolation_source="patient with severe acute respiratory
syndrome"
/db_xref="taxon:230471"
/country="China: Hangzhou"
1..6
/note="transcription regulatory core sequence 6"
misc_feature
162..353
/codon_start=1
/product="uncharacterized protein 6"
/protein_id="AAP27473.1"
/db_xref="GI:38735510"
/translation="MFHLVDQVTVIAELIIMRTFRIALWLDVIISIVRLQFLKL
TKKNYSELDDSEPMELDYP"
331
/note="Xiaohong genotype locus 6"
misc_feature 355..360
/note="transcription regulatory core sequence 7"
361..729
/codon_start=1
/product="uncharacterized protein 7a"
/protein_id="AAP27474.1"
/db_xref="GI:38735511"
/translation="MKILFLTLVFTSCELYHQVCVRGTTVLLKEPCPSGTGNS
PFHPLADNKFALTCTCTHFAFACADGTRHTYQLRARSVSPKLPFIRQEEVQQLYSPLF
LIVAAVLFLICFTIKRTE"
726..860
/codon_start=1
/product="uncharacterized protein 7b"
```

```
/protein_id="AAR27475.1"
/db_xref="GI:38735512"
/translation="MNELTLIDFVLCFLAFLFLVLMLIFWFSLEIQDLBECTVK
"
misc_feature
861..866
/feature="transcription regulatory core sequence 8"
CDS
867..986
/codon_start=1
/product="uncharacterized protein 8a"
/db_xref="AAR27476.1"
/db_xref="GI:38735513"
/translation="MKLLIVLTCISLSCICTVQRCASNKPVLDEDPCKVQH"
misc_feature
915
/feature="Xiaohong genotype locus 7"
952..1206
/codon_start=1
/product="uncharacterized protein 8b"
/db_xref="AAR27477.1"
/db_xref="GI:38735514"
/translation="MCKLILVRYNTRGNTYSTAWLCAIGKVLPHRWHTMVQCTPNU
TINCQDPAGGALIAFCWYLHEGHTAANFRDLVLVNLKRTN"
misc_feature
1193..1198
/feature="transcription regulatory core sequence 9"
1208..2476
/codon_start=1
/product="nucleocapsid protein"
/db_xref="AAR27478.1"
/db_xref="GI:31505970"
/translation="MSDNGPQSNORSAPRITFGGPTDSTNNONGRNGARPKORRQK
GLPNTASWFTALTOGKEELFRPGQVPLNTNSGPDQIGYVRATRRVGGDGKM
KELSPRWTFYILGTPEASLPYKANKEGIVWATEGALNPKDHIGTRPNNAATVL
QLPQGTLLPKGAEBSGGSGQSSRSRSGNSRSTPGSSRSGNSPQSNVATVL
ALALLLNLKSVGSGQQGQQTWTKSAEASKPRKRTATQNVNTOAFG
RGPPTQCNFGDQLIRGQTDYKHPQIQAPASAFPGMSRIGMEVTPSGTWLTY
HGAIKLDDKDPQFQKVNILNKHIDAYKTFPPTPEPKDKKKKTDEAQLPQKQKQPT
VTLLPAADMDPSRQLQNSMSGASADSTQA"
1218..1514
/codon_start=1
/product="uncharacterized protein 9b"
/db_xref="AAP69659.1"
/db_xref="GI:31747858"
/translation="WDPQNTVVPVPPALHLVDPIQILITIMEDAMGQGNADPKVYP
ITIRLGQSLSMARNLDSLEARAFQSTPIVVQMTKLATTEELPDEFVVTAK"
1671..1883
/codon_start=1
/product="uncharacterized protein 9c"
/db_xref="AAP69660.1"
/db_xref="GI:31747859"
/translation="MLPPCYNFLKQHCQKQASTQREAEANVKPLLAPHVAVIQEIQ
LLAAVGEILLLEWLAEVVKLPSRYCC"
ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 2810;
Best Local Similarity 100.0%; Pred. No. 0.033; 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

Qy 1 GCTGTGAACCAAGCAGCAGTATTAT 25
|||||
Db 1374 GCTGTGAACCAAGCAGCAGTATTAT 1350

RESULT 6
AY443086S10/c
LOCUS
DEFINITION
SARS coronavirus CUHK-L2 orf4 and orf3 genes, partial cds; small
envelope E protein, membrane glycoprotein M, orf7, orf8, orf9,
orf10, and orf11 genes, complete cds; nucleocapsid protein gene,
partial cds; and orf13 gene, complete cds.
ACCESSION
AY443095
VERSION
AY443095.1
KEYWORDS
10 of 10
SEGMENT
SARS coronavirus CUHK-L2
```

```
ORGANISM
SARS coronavirus CUHK-L2
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
1 (bases 1 to 3080)
AUTHORS
Chim,S.C., Tong,Y.K., Hung,E.C.W., Chiu,R.W.K. and Lo,Y.M.D.
Genomic Sequencing of a SARS Coronavirus Isolate That Predated the
Metropole Hotel Case Cluster in Hong Kong
Clin. Chem. 50 (1), 231-233 (2004)
JOURNAL
14709660
PUBMED
2 (bases 1 to 3080)
AUTHORS
Chim,S.C., Tong,Y.K., Hung,E.C., Chiu,R.W. and Lo,Y.M.
Direct Submission
Submitted (20-OCT-2003) Chemical Pathology, Chinese University of
Hong Kong, Prince of Wales Hospital, 30-32 Ngan Shing Street,
Shatin, Hong Kong, China
FEATURES
Location/Qualifiers
1..3080
/organism="SARS coronavirus CUHK-L2"
/virus
/mol_type="genomic RNA"
/db_xref="taxon:260550"
<1..328
/codon_start=2
/product="orf4"
/protein_id="AAS01066.1"
/db_xref="GI:41352897"
/translation="VVILRIQVLKTMSLYMAISPFTTSLSLHKLQTLVLKMLHS
SSLTSLKLTHERMCKYTOSTALQELLIOOWIFMMSRRLLACLCKHKVKVSTNLCTHSF
RKQVR"
<1..267
/codon_start=1
/product="orf3"
/protein_id="AAS01065.1"
/db_xref="GI:41352896"
/translation="GYSSEDRHSGYKDYVVVHGYFTEVYQLESTQITTTDTGIE NATF
FIENKLVKDPNPVQIHTIDGSSGVANPAMDPIYDEPTTTTISVPL"
292..522
/codon_start=1
/product="small envelope E protein"
/protein_id="AAS01067.1"
/db_xref="GI:41352898"
/translation="MYSFVSEETGLIVNSVLLFLAFVVFVLLVTLAILTALRLCAYCC
NIVNVSIVKPIVYVYVRKYNLSSEGVFDLLV"
573..1238
/codon_start=1
/product="membrane glycoprotein M"
/protein_id="AAS01068.1"
/db_xref="GI:41352899"
/translation="WADNGTITVEELKOLLQEWNLVIGFLAWIMLLQFAYSNRNR
LYIKLVFLWLLWPTLACFLVLAAYRINWVTGGIACIACIAGIWMLSYFVASRFLF
ARTRSMFSNPETNILLNVPKRTIIVTRPMESELVIGAVIIRGHLRMAGHSLGRCDI
KDLPKETVATSRTLSYVYKLGASQVRGTDSGFAAYNRVRYRIGNYKLTNDHAGSNDNIAL
LVQ"
1249..1440
/codon_start=1
/product="orf7"
/protein_id="AAS01069.1"
/db_xref="GI:41352900"
/translation="MFHLVDQVTVVAEILIIIMRTFRIAMNLDVLISSIVRLFKPL
TKKNVSELDDEPMELDYP"
1448..1816
/codon_start=1
/product="orf8"
/protein_id="AAS01070.1"
/db_xref="GI:41352901"
/translation="MKIILFLTLIVFTSCELYHQECVGRGTTVLKKEPCPSGTGEGNS
PFPPLADNKEALTCTSTHFACADGTRHTYQLRARSVPKLFIRQEEVQQLYSPLF
LIVAAVLFLICLFTIKRTE"
1813..1947
/codon_start=1
/product="orf9"
/protein_id="AAS01071.1"
/db_xref="GI:41352902"
```



```
/translation="MNELTLIDFVLCFLAFLLFLVLIMLIIFWFSLEIQDLBPCTKV
"
1954. .2073
/codon_start=1
/product="orf10"
/protein_id="AA01072.1"
/db_xref="GI:41352903"
/translation="MKLLIVLTICLSLCIRTVVQRCASNKPHELDPCVKQH"
2039. .2293
/codon_start=1
/product="orf11"
/protein_id="AA01073.1"
/db_xref="GI:41352904"
/translation="MCLKILVRYNTRGNTYSTAWLCALGKVLFPFHRMHTVQICTPNV
TINCQDPAGGALLARCWYLHEGHQTRAFRDVLVLNKRIN"
2295. .3080
/codon_start=1
/product="nucleocapsid protein"
/protein_id="AA01074.1"
/db_xref="GI:41352905"
/translation="MSDNGPQSNQSRAPRTIFGPTDSTDNNGRNGRNGARPKORPQ
GLPNTASFTALTQHGKELRPRGQGVPIINTSGPDQIGYRTRATRVRGDGKM
KELSPWFYFLGTGPEASLPYGANKEGIIVWATEXXXXXXXKXXXXXXXKX
XLPGQTLTPKFGAEGSRGGSQASRSRXXNSRSTPGSSRGNSPARMASGGGET
ALALLLDRLNQLSVKSGKQQQGXVTYTKSAEASKKPRQK"
2305. .2601
/codon_start=1
/product="orf13"
/protein_id="AA01075.1"
/db_xref="GI:41352906"
/translation="MPNQNVVPPALHLVDPOQLTIEMEDAMGQONSADPKVYP
IILRLGSQLSWARRNLSLEARFQSTPIVVQMTKLATTELPDEFVVVYAK"

ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 3080;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCAGCAGTATTAT 25
|||||
Db 2461 GCTGTGAACCAAGCAGCAGTATTAT 2437

RESULT 7
AY304489/c
LOCUS 8439 bp RNA linear VRL 05-NOV-2003
DEFINITION SARS coronavirus SZ1, partial genome.
ACCESSION AY304489
VERSION AY304489.1 GI:34482140
KEYWORDS SARS coronavirus SZ1
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 8439)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 12958366
PUBMED
REFERENCE 2 (bases 1 to 8439)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
Location/Qualifiers
source 1. .8439
/organism="SARS coronavirus SZ1"
/mol_type="genomic RNA"
/isolate="SZ1"
/db_xref="taxon:231514"
/country="Hong Kong"

ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 8581;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCAGCAGTATTAT 25
|||||
Db 7150 GCTGTGAACCAAGCAGCAGTATTAT 7126

RESULT 9
AY304491/c
LOCUS 11006 bp RNA linear VRL 05-NOV-2003
DEFINITION SARS coronavirus GZ60, partial genome.
ACCESSION AY304491
VERSION AY304491.1 GI:34482142
KEYWORDS SARS coronavirus GZ60
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 11006)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
```

```
/mol_type="genomic RNA"
/isolate="SZ1"
/db_xref="taxon:231516"
/country="Hong Kong"

ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 8439;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCAGCAGTATTAT 25
|||||
Db 7156 GCTGTGAACCAAGCAGCAGTATTAT 7132

RESULT 8
AY304487/c
LOCUS 8581 bp RNA linear VRL 05-NOV-2003
DEFINITION SARS coronavirus SZ13, partial genome.
ACCESSION AY304487
VERSION AY304487.1 GI:34482138
KEYWORDS SARS coronavirus SZ13
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 8581)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 12958366
PUBMED
REFERENCE 2 (bases 1 to 8581)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
Location/Qualifiers
source 1. .8581
/organism="SARS coronavirus SZ13"
/mol_type="genomic RNA"
/isolate="SZ13"
/db_xref="taxon:231514"
/country="Hong Kong"

ORIGIN
Query Match 100.0%; Score 25; DB 14; Length 8581;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCAGCAGTATTAT 25
|||||
Db 7150 GCTGTGAACCAAGCAGCAGTATTAT 7126

RESULT 9
AY304491/c
LOCUS 11006 bp RNA linear VRL 05-NOV-2003
DEFINITION SARS coronavirus GZ60, partial genome.
ACCESSION AY304491
VERSION AY304491.1 GI:34482142
KEYWORDS SARS coronavirus GZ60
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 11006)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
```

Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,  
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and  
Poon,L.L.M.

TITLE Isolation and characterization of viruses related to the SARS  
coronavirus from animals in southern China

JOURNAL Science 302 (5643), 276-278 (2003)

MEDLINE 22913660

PUBMED 12958366

REFERENCE 2 (bases 1 to 11006)

AUTHORS Guan,Y. and Zheng,B.J.

TITLE Direct Submission

JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,  
University Pathology Building, Queen Mary Hospital, Pokfulam Road,  
Hong Kong, China

FEATURES Location/Qualifiers

1..11006  
/organism="SARS coronavirus GZ60"  
/mol\_type="genomic RNA"  
/isolate="GZ60"  
/db\_xref="taxon:231518"  
/country="Hong Kong"

ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 11006;

Best Local Similarity 100.0%; Pred. No. 0.033;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCGCAGTATTAT 25

|||||

Db 9557 GCTGTGAACCAAGCGCAGTATTAT 9533

RESULT 10

AY304493/c

LOCUS AY304493

DEFINITION SARS coronavirus HKU-65806, partial genome.

ACCESSION AY304493

VERSION AY304493.1 GI:34482144

KEYWORDS

SOURCE SARS coronavirus HKU-65806

ORGANISM SARS coronavirus HKU-65806

Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE 1 (bases 1 to 11010)

AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,  
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,  
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and  
Poon,L.L.M.

TITLE Isolation and characterization of viruses related to the SARS

JOURNAL Science 302 (5643), 276-278 (2003)

MEDLINE 22913660

PUBMED 12958366

REFERENCE 2 (bases 1 to 11010)

AUTHORS Guan,Y. and Zheng,B.J.

TITLE Direct Submission

JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,  
University Pathology Building, Queen Mary Hospital, Pokfulam Road,  
Hong Kong, China

FEATURES

Location/Qualifiers  
1..11010  
/organism="SARS coronavirus HKU-65806"  
/mol\_type="genomic RNA"  
/isolate="HKU-65806"  
/db\_xref="taxon:231520"  
/country="Hong Kong"

ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 11010;

Best Local Similarity 100.0%; Pred. No. 0.033;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCGCAGTATTAT 25

Db 9561 GCTGTGAACCAAGCGCAGTATTAT 9537

RESULT 11

AY304494/c

LOCUS AY304494

DEFINITION SARS coronavirus HKU-66078, partial genome.

ACCESSION AY304494

VERSION AY304494.1 GI:34482145

KEYWORDS

SOURCE SARS coronavirus HKU-66078

ORGANISM SARS coronavirus HKU-66078

Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE 1 (bases 1 to 11010)

AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,  
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,  
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and  
Poon,L.L.M.

TITLE Isolation and characterization of viruses related to the SARS

JOURNAL Science 302 (5643), 276-278 (2003)

MEDLINE 22913660

PUBMED 12958366

REFERENCE 2 (bases 1 to 11010)

AUTHORS Guan,Y. and Zheng,B.J.

TITLE Direct Submission

JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,  
University Pathology Building, Queen Mary Hospital, Pokfulam Road,  
Hong Kong, China

FEATURES

Location/Qualifiers  
1..11010  
/organism="SARS coronavirus HKU-66078"  
/mol\_type="genomic RNA"  
/isolate="HKU-66078"  
/db\_xref="taxon:231521"  
/country="Hong Kong"

ORIGIN

Query Match 100.0%; Score 25; DB 14; Length 11010;

Best Local Similarity 100.0%; Pred. No. 0.033;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCGCAGTATTAT 25

|||||

Db 9561 GCTGTGAACCAAGCGCAGTATTAT 9537

RESULT 12

AY304490/c

LOCUS AY304490

DEFINITION SARS coronavirus GZ43, partial genome.

ACCESSION AY304490

VERSION AY304490.1 GI:34482141

KEYWORDS

SOURCE SARS coronavirus GZ43

ORGANISM SARS coronavirus GZ43

Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE 1 (bases 1 to 13471)

AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,  
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,  
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and  
Poon,L.L.M.

TITLE Isolation and characterization of viruses related to the SARS

JOURNAL Science 302 (5643), 276-278 (2003)

MEDLINE 22913660

PUBMED 12958366

REFERENCE 2 (bases 1 to 13471)

AUTHORS Guan,Y. and Zheng,B.J.

TITLE Direct Submission

JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Pokfulam Road, Hong Kong, China

FEATURES source  
Location/Qualifiers  
1. .13471  
/organism="SARS coronavirus GZ43"  
/mol\_type="genomic RNA"  
/isolate="GZ43"  
/db\_xref="taxon:231517"  
/country="Hong Kong"

ORIGIN  
Query Match 100.0%; Score 25; DB 14; Length 13471;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTGTGAACCAAGCAGCAGTATTAT 25  
|||||  
Db 12022 GCTGTGAACCAAGCAGCAGTATTAT 11998  
|||||

RESULT 13  
AY304492/c  
LOCUS AY304492 13471 bp RNA linear VRL 05-NOV-2003  
DEFINITION SARS coronavirus HKU-36871, partial genome.  
ACCESSION AY304492  
VERSION AY304492.1 GI:34482143  
KEYWORDS  
SOURCE  
ORGANISM SARS coronavirus HKU-36871  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.  
REFERENCE 1 (bases 1 to 13471)  
AUTHORS Guan, Y., Zheng, B. J., He, Y. Q., Liu, X. L., Zhuang, Z. X., Cheung, C. L., Luo, S. W., Li, P. H., Zhang, L. J., Guan, Y. J., Butt, K. M., Wong, K. L., Chan, K. W., Lim, W., Shortridge, K. F., Yuen, K. Y., Peiris, J. S. M. and Poon, L. L. M.

TITLE Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China  
JOURNAL Science 302 (5643), 276-278 (2003)  
MEDLINE 22913660  
PUBMED 12958366  
REFERENCE 2 (bases 1 to 13471)  
AUTHORS Guan, Y. and Zheng, B. J.  
TITLE Direct Submission  
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Pokfulam Road, Hong Kong, China

FEATURES source  
Location/Qualifiers  
1. .13471  
/organism="SARS coronavirus HKU-36871"  
/mol\_type="genomic RNA"  
/isolate="HKU-36871"  
/db\_xref="taxon:231519"  
/country="Hong Kong"

ORIGIN  
Query Match 100.0%; Score 25; DB 14; Length 13471;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTGTGAACCAAGCAGCAGTATTAT 25  
|||||  
Db 12022 GCTGTGAACCAAGCAGCAGTATTAT 11998  
|||||

RESULT 14  
AY463060/c  
LOCUS AY463060 29013 bp RNA linear VRL 05-JAN-2004  
DEFINITION SARS coronavirus ShanghaiQXC2, complete genome.  
ACCESSION AY463060  
VERSION AY463060.1 GI:40457448  
KEYWORDS

SOURCE  
ORGANISM

SARS coronavirus ShanghaiQXC2  
SARS coronavirus ShanghaiQXC2  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 29013)  
Yuan, Z., Zhang, X., Hu, Y., Lan, S., Zhou, Z., Wang, H. and Wen, Y.  
Analysis of SARS coronavirus genome in Shanghai isolates  
Unpublished  
2 (bases 1 to 29013)  
Yuan, Z., Zhang, X., Hu, Y., Lan, S., Zhou, Z., Wang, H. and Wen, Y.  
Direct Submission  
Submitted (11-NOV-2003) Key Lab of Medical Molecular Virology,  
Shanghai Medical College, Fudan University, 138 Yi Xue Yuan Road,  
Shanghai 200032, P.R. China

FEATURES  
source

1. .29013  
/organism="SARS coronavirus ShanghaiQXC2"  
/viralion  
/mol\_type="genomic RNA"  
/strain="ShanghaiQXC2"  
/db\_xref="taxon:258508"  
join(186. .12734, 12734. .20827)  
/notes="ribosomal slippage"  
/codon\_start=1  
/product="orf1ab"  
/protein\_id="AAR86774.1"  
/db\_xref="GI:40457449"

## CDS

/translations="MESLVGVNEKTHVQLSLPVLQVRLVRFVGFQDSVEALSEARE  
HLKNGTCGLVELEKGLPOLEQPVYFKIRSDALSTNHGHVKVFLVAGMDGIQYGRSGI  
TLGVVPHVGETPIAYRVLLRNGKNGAGHSYGLDKSYDLGDELGTPIEDYEON  
WNTKSGALRELTRELNGAVTRYVDNFCGPDGYPDCIKDFLARAGSMCTLSEQ  
LDYIESRGVYCCRDHEIAFTERSDYSYHETPEIKSAKKDFPKGCEPKFVP  
LNSKVIOPRVEKKETLIEGTTCGLPTNAVWMPACQDPEIGKHSVADYHNHS  
DFLKATCEHCCTENLVIEGTTCGLPTNAVWMPACQDPEIGKHSVADYHNHS  
IETLRKGRTRCFGCVFAYGVCKYKRAVWPRASADIGSGHTGIGDNVETLNEEDL  
LEILSRVNNIVGDFHLENEVAIILASFSASTSARIDIKSLDYKSPYITVESCRA  
YKTRKPKVGNWICQGRSLVPLCGFPQQAQVIRSIPTARDLAAHNSIPDQRA  
VITLDIGISQSLVADWYVTSDDLNTNSVIMAYVTGGLVQQTQSWLSNLGTTVEKL  
RPIFEWEAKLSAGVFLKDAWEILKITGVFDIVKQIQVADSNIKDCVCYKIDV  
NKALEMIDQVITAGAKLSNLGEVFIQAKGLYRQCTIRGKQLOQLMPLPKPEVT  
FLGSDSDVLTSTFVNLKNGELEALETPVDSFTNGAIYCTPVCVNGLMLEIKDKSQ  
YCALSPGLATNNVRLKGPAPIKVTFGEDTWEVQGYKRVITFELDRVVKVNE  
KCSYVTSGETVEFVACVAAVKTLPQVSLDNLNMGIDLDDEWSATFLPDDAGE  
ENFSRMYSFPDDEEDDAECBEEIEDTCEHEYGTEDDYGLGFLFGASAEYTR  
VEEBEEDLDDTTEQSEIPEPEPTPEEPVNOFTGYLTDNVAIKCVDTVLEAQA  
NPWIVNAANIHLKGGVAGALNKATNGAMOXESDDYIKNGPLTVGGSCLLSGHNL  
AKCLHVGNLNAGEDIOQLKAAYENFSQDILLAPLSAGIFGAKPLQSLQVOT  
VRTQVIAVNDKALYEQVMDYLDNLKPRVEAKQEPPEPNTSEKTSKSVQKPVV  
KPKICACIDEVITTLBETPLTNKLLFADINKGYHDSQNLRGDMSFLEEDAPYM  
VGDVITSGDITCVIIPSKAGGTTEMLSRALKKVPVDEYITTYPGQACAGTYLEAKT  
ALKKCSAFYVLPSEAPNAKEELGTVSWNLREMLAHAEAREAKLMPICMDVRAIMAT  
ORKYKGVKIQEGIVDYVRFPPFVTSKEPVASIIITKLNSLNEPLVTMPIGVYTHGFNLE  
EAAKCRSLKAPAVSVSPDVAITYNGYLTSSKTSSEHFVETSLAGSYRWMSYS  
QRTGELGVEFLKRGDKIYHTLESFVFLHDEVLSDKLKSLSLRVRKVIKVTFTVD  
YNTLTLQVDMSTYGGQFTYLDGADVTKIKPHVNHEGKTFVLPDSDTLRSEAFE  
YHTLDESFLGRYMSALNHTKWKFTSPQVGLTSIKWADNNCYLSVLLALQQLSEAF  
APALQEAAYRARGADAACALILAYSNKTVGELGVRETMTHLLOHANEASARVLN  
VCKHCGKQTTTLTGVAVMVMTGLSLVDNLKTVGSI PCVGRDATHLVQESSFVMM  
SAPPAGYKQQGTFLCANEYTGNYQGHYTHITAKETLYRIDGNHLTKMEYKGPVTD  
VFKYETSYHTTIKPVSKLDGTYLEBPKLDGYKKNAYVENTSIITKKNPNSLSA  
LGLTIAHTIAAINSVPMSKILAYVYKFLGQAAITTSNCAKRLAQRFVMPYVFT  
LILQLCTGTRNSIRASPLTIKNSVKSVAKLCLDAGINTVYKSPFKLFTIAMW  
LILLSICLGLSLICVTAAGVLLSNFAPSYCNGVRELYLNSNVTTMDFCBGSFPCS  
CLSGDLSDYPALETIOVLTISYKLDLTILGLAAEWLAVMLFTFVYLLGSAIMQ  
AFPGYPASHFISNSWLMWFIISVQAPVSAVMYIIFASFYIYKSYHIMDGCTIS  
STCMCYKCNHATRVECTIIVNGKRSFYVYANGRGFGFKTHNWNCLNCDTFCGTSTF  
ISDEVARDLSQKRPINPTDQSSYIVDSVAVNGALHPYFDKAGQTYRHPHLSHFV  
NLIDNRANNTKGLRPINPTDQSSYIVDSVAVNGALHPYFDKAGQTYRHPHLSHFV  
GDSYEVSMDFADVDTFSAPVMEKALVATAHSELAKGVALDGLVSTFVSAR  
QGVVDVDTDKVIECLKLSHSDLEVTGDCSYNFMVYKKNVENMPTDRIAGACIDNA  
RHINAQVAKSHVNSLIIWNKDYMSLSQLRKQIRSAKQNNIIFRLTCATIRQVWVI  
TTKISLKGKGIKSTCFKMLKLTLLCVLAALVCYI VMPVHTLSHDGYTNEIGYKAI  
QDGVTRDIIISTDTCFANKAGDFAWFSQSGSYKNDKSCPVAAIITREITGFIYVGLP

GTVALRANGDFLHFLPVSAGVNICYTPSKLIEYSDPATSAVLAAECTIFPKAMGK  
PVPYCYDNLEGSISYSELRPDTRYVLMDGSIIOFPNTYLEGSRVYVTTTDEAECRH  
GTCSESEVICLSTSGRWNNHNEHRALSGVFCGVGDVADNLIANIIFTLVQVPGALDAS  
ASVAGGAIILATVCAAYFMKFRFAEGEYHNHVAANALLFLMSTLCLVPAYSLP  
GVYSYFNGVTFSTFEEAALCTFLNKKEMYLKLRSETLLPQYNYRLALYNKKYFS  
LKRKRYMNGVTFSTFEEAALCTFLNKKEMYLKLRSETLLPQYNYRLALYNKKYFS  
GVLDTSVREAAKCALANDSNGADVLPQPTQSTSAVGLDLPKMAFSPGKV  
EGCVQVTCGTTNGLMDLVYCPRHVICTAEDMLNPNYEDLLIRKSNHFLVOAG  
NVLQRIHSGMNCRLRUKVDISNPKYKFKVRIQPGQTFSLVACNGSPSGVYQCA  
MRPNHTIKGSFLNGSGVFNIDYDVSFCYMHHEMLPTGVHAGTDLEGKPYGFVD  
RTOAAGVETITLNLAWLYAAVINGDRWFLNFTTLLANDFNLVAMKYNPELTD  
HYDILGPLSAOQIAVLMDCAALKELLONGMNGRTILGSLTLEDEFTFDVVROCSG  
TFQGEKPIKGTGHEHMLLTPLTSLILVQSTOWSLFFEVVENAFLPTLPLMAAAC  
AMLLVKHAEFLCLFLPSLATVAFNMYMPASWVRIMTWLEADTSLSGYRLKDC  
VNYASGVLLITLTVDDAARRVYLLMNVITLVYKVVYGNALDQALSMWALVISA  
TSNYGCVETIMFLARAVFVCEVYPLLFTIGNTLQCMILVYCFGLCCCCYGLFC  
LNNRYFRITLGVYDLVSTQEPFRYNSQGLPKPSIDAFKLNLILGILGGKPKICVA  
TVQSKMSDVKTSVLLSVLOQLRVSESSKLWAOCVQHLNDILLAKDTTEAFKMWSL  
LSVLLSMOGAVDINRLCEVLDNRATLOALASEFSLPSYAAYATAQAYCAVANGD  
SEVLKULKLSNVAKSEPRDAAMQRLERKADQMTQMYKQARSEKRAKYSAMQ  
TMFTMLRLKLDNLANNIINNADGCVPLNIIPLTAAKLMVVYDGYKNTCDGNT  
FYTAALMEIOQVVDADSKI VQSEINNDNSENLAWPLITVLRANSAAKIQNNELSP  
VALRQMSCAAGTQTACTDNALAYNNSKGRFVLLALLSHQDLKWARFPKSDGTG  
IYTELEPCRCFVYDTPKPKMYLYPIKGLNLRNGMVLGSLAATVRLQAGNATVPA  
NSTVSECAFAVPAKAYKVLASGGQPLTNCVKMLCTHGTGQATVRLPEANMDQS  
FGAASCLLYCRCHIDHPNPKGCDLUGKIVQIPTTCANDPVGFTLRNVTVCGMWKG  
YGCSCDLREPLMGADASTFFKRGVSAARLTTCGTGTSTVDVYRAPFIYNEKVG  
FAKFLTKNCFQEKDEGNLLDSYFVKRHTMSNYOHEETIYNLIKDCPAVAYHDF  
DMYFVEMPDILRVYANLGERVQSLTKVQCDAMRDAGIVGLVTLDNQDLNGMYD  
PGDFVQVAGCGVPIVDSYSLMLPILITLALAAESHMDADLAKPLIKWLLKYDFT  
BERCLFDRYFKWQOTHPNCINGCLDRCLHCANFNVLSTVFPPTSPGRLVKRIF  
VDBGVFPVSTYGHFRELGVHNDVNLHSSRLFKELIYAADPAMHAASGNLLDKR  
TTCFYVALTNLPMDICRTVPKPGNFDFYDAVSKGFKEGSEVELKHPFFAQDGNAAI  
STSYRYNLNPMTCIDRQLLFVVEVDYFCYDCGCGNANQVIVNNLDKAGGFENK  
WCKARLYVDSMSVEDQDALFAYTKRNVIPITOMNLKVAISAKNRARTVAGVSICTM  
"NRQPHOKLISATRATVIGITKEFGGWHMLKTVYSVDETHLMDYDKCDR  
AMPNMLRIMASLVLARKNNTCCNLSHRFYRLANECAQVLSNMGCGSLYKPGGTS  
GDATTAYANSFNI COAVTANVALLSTDGNKIDKYVRNIQHRLEYCLYRNRVDHE  
FYDEFYALRKHSMILLSDDAVVCYNSYAAQGLVASIKNFKAVLYQNNYFMSEAK  
CWTELDLGHFECQHTMLVKQGDYVLYPDPSPRILGAGCFVDVDTIKTQDGLMI  
BREFSLADAYPLTKHPNOEYADVPHLYQYIRKLHDELTHGLMDVMSVMTNDNTSR  
YHEPEYEMVTPHTVLQAVGACVLCNSQTSRLCGACIRRPFLCCKCCYDHYIYSHK  
LVYVNPVYAPAGDQTDVDTQIYLGMSIYCKSHKPPILSPFLCANGQVFGYKNTCV  
GSDNVTFNAIATCDWNTAGDYILANTCTERLKHFAAETLKATETFKLSYGIATVRE  
TYYKLNVDGYFVLTSHTVMPLSAPTLPOEHYVRIITGLYPLTNIISDEFSSNVANYQV  
SMKYFGIPGERTCCLDKRATCTSSDTYACNHSVGFDYVYVPMFDVQWQGTG  
NLIQNSHDQCHQVGHNAHVASCDAIMTRCLAVHECFVRVDSVEYPIIIGDELVRNSAC  
RKQVHMVKSALLADKFPVLHIGNPKAIKCVPAQVEMWPKYDAQPCSDKAYKIBELF  
VSYATHDKFTDGVCLFNNCNDVRYPANALVCRFDTVRVLSNLPSCDGGSLYVNHGA  
FYTPAFDSAPTNLQQLPPFYSDSPCESHGQVYVSDIDVYPLKSACTNLCNLGAV  
CHHANEAFYLDANMMISAGFSYKQFDYTNLNDITLQSLSENVANVNVNKH  
FDGHAAGEAPVITINNAVYTKVDGIDVEITPESTKTLFPNVAFELMAKRNIPKPEIKIL  
NNLGVDIANTVIMDYKREAPASVTIGVCTMDIAXKPTESACSLTVLFDPRVGG  
VLFPNARNGLVITGESSVKGLTPSKPAQASVNGVTLIGESVKTOFNTPKVDGIIOQ  
LPETVYQSDRLDEPKPRSMETDPLELAMDDEFIORYKLEGYAFEHIVYDFPSHQGL  
GLHLMGLAKRQSDPLKLEDPIMDSVTNVIYFIDTAQIGSSKCVSRVIDLDDDEVE  
LTKSDQLSVKVKVYITDIYAESIFMLCKOGHVEITFPLOASQAWQPGVAMPNLYK  
MQRMLEKCDLQNGENAVIPKIMMNVAKYTQLCQYLNLTILFLAVYNNRVHFGAGS  
DKGAPGTVLAVLPTGLTLLDNDLNDVSDASTLITGDCAVTHANKWDLIISDMY  
DRPKHVTKENDSKGFFTYLGCFLIKQKALGDSIAVKITESHMADLYKLMGFHSW  
TAPVTNVNASSSEAFIIGANYLKGPKQIDGYTMHANYI FWRNTNPIOLSSYSLFDSM  
KFLPLKRGTVAMSLKENQINDMISLLEKGRILIRENNRVVYVSSDILVNN"

20834. .24601  
/codon\_start=1  
/product="spike"  
/protein\_id="AAR86775.1"  
/db\_xref="GI:40457450"  
/translation="MFIPLFLTLTSGSDLLDCTFTDDVQAPNTQHTSSMRGVYYPD  
EIGSPNDLYLQDLPLFPYSNVTGPIINHFGNVPVFPKDGIVFAATKSNVVRGW  
EIGSPNNKSSQSVIINNSTNVVIRACNPELNDNFEAVSKPMGTQTHIMI FDNA  
FISYSDAFSLDSEKSGNFKHLRFVFNKNGDPLVYKGYQPIDVRLDPGFWTLKP  
IFKPLGINITNFRAILTAFSPAQIDWTSAAAFVGYKLPITFMKTDENGITIDAV  
DCSQNPLAEIKCSVKSEIFDKGIYQTSNFRVPSGDVVRFPNITNLCPFGVFNATKF  
PSVYAMRKILSNCAVDSVLYSTFFSTFKCYGVSATKNDLFCFNYSYKRLHSGKLRP  
DVOIAPGOTGVIAVYKLPDDPMFCVLANWTRNIDATSTGNVYKRYLHSGKLRP  
FERDISNVPSDPCKPTCPALNCYVPLNDGYFTTGTGIGYQYVRVVLSELNABA  
TVCGPKLSTDIKNCQNFNFNGLTGCVLPSSKRFPQPPQGRDVSDDTSVDRDPK  
TSEILDSPCAFGGVSVITPGTNASSEVLYQDVNCTNVSAAHADQUTPAWRIYST  
GNNYFOAQGLIGAEHVEYSEDIPIGAGICASYHTVLSLLRSTQSKSIYATVMSLG  
ADSSIAYSNNTIAIPTNFSITITTEVMPVSMKTSVDCNMVTCGSTECANLLQYGS  
PCTOLNALSIGIAEODRNTREVAQVKMYKPTLKYFGGFNFSOILPDLPLKPKRS  
PIELLFNKVTLADAGPMKQYGECLGDINARDLLCAQKENGITVLPPLTDDMLAAVT  
AALVSGTATAGTFCAGAAQIIPFAMQAYRPNIGVIGVTONVLYENKQIQAQFNKAIS  
QIQBSLITSTALGKLQDVNQNAQALNTLVQLSSNFGAISVNLNDILSRDKVEAE  
VQIDRLITGRLOSQTQYVTOQIIRAAETIRASANLAATKMSCVLQSKSRVDFCGGYH  
LMSFPAQHGWFVHVTYVPSQBRNFTTAPACHEKAYFPREGVFNENGTSWFTIQ  
RNFPSPQIITDNTFVSGNCDVIGIINNTVVDLOPELDSFGELDKVYFNKHTSPDV  
DLGDISINASVNIQKIDRLNEVAKNLSLIDLQELKYEQIKWPNVWLGFLIA  
GLIAIWMVITLLCCWTCSCCLKGCACSCGCKFDEDDSEPVKGVKLUHYT"  
24610. .25434  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAR86776.1"  
/db\_xref="GI:40457451"  
/translation="MDLPMRFFTLGSITAPQVPIKIDNASPASTAHATATIPLOASLPFG  
WLTVGAPLAVFQATKIIALNKKWQALYKGFQFINLLLFTVTLVSHLLVLAAGME  
AQFLYLXALYFQACINAKIRMCWCKCKSNPLLYDANYFCWHTNYDYCIPIY  
NSVTDITVTEGDDGSTPKLXEDYQIGGSDRHSQVGVVYVHGYPTEVYVQLSESTQ  
ITNTGTIENATFEIFNKLVKDPNVQIHTIDGSSGVANPAMPDIYDEPTTTTSLVPL"  
25031. .25495  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAR86777.1"  
/db\_xref="GI:40457452"  
/translation="MMPITLPAFTHTMTVTHIVTSOIQLSLLKATAFQHNSKKT  
KLVLRLIGTQVILKMSLYMALSPEFTTSLSHKLQTLVLKMLHSSSLTSLKTHRM  
CKYQSTALQKLLQLQWLIQFMRRRLACLCKHKBSINLCTHSTRKKQVR"  
25459. .25689  
/codon\_start=1  
/product="E"  
/protein\_id="AAR86778.1"  
Query Match 100.0%; Score 25; DB 14; Length 29013;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCTGTGAACCAAGCGCAGTATTAT 25  
DB 27628 GCTGTGAACCAAGCGCAGTATTAT 27604  
RESULT 15  
LOCUS AY394999 29350 bp RNA linear VPL 29-JAN-2004  
DEFINITION SARS coronavirus LC2, complete genome.  
ACCESSION AY394999  
VERSION AY394999.1 GI:37624342  
KEYWORDS SARS coronavirus LC2  
SOURCE SARS coronavirus LC2  
ORGANISM SARS coronavirus LC2  
REFERENCE Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
AUTHORS Coronaviridae; Coronavirus.  
1 (bases 1 to 29350)  
CONSTRM The SARS epidemiology consortium of Guangdong

Search completed: May 16, 2005, 02:00:25  
Job time : 597.656 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 15, 2005, 17:17:00 ; Search time 320.703 seconds  
(without alignments)  
461.466 Million cell updates/sec

Title: US-10-808-187A-2475  
Perfect score: 25  
Sequence: 1 gctgtgaaccaagacgagctattat 25

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N Geneseqn\_16Dec04:\*
- 1: Geneseqn1980s:\*
  - 2: Geneseqn1990s:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002as:\*
  - 7: Geneseqn2002bs:\*
  - 8: Geneseqn2003as:\*
  - 9: Geneseqn2003bs:\*
  - 10: Geneseqn2003cs:\*
  - 11: Geneseqn2003ds:\*
  - 12: Geneseqn2004as:\*
  - 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25	100.0	29751	12 ADJ39000	SARS coro
2	17.8	71.2	1533	12 ADM74528	Adm74528 thale cre
3	17.8	71.2	1557	3 AAC49448	Aac49448 Arabidops
4	17.8	71.2	1792	3 AAC49462	Aac49462 Arabidops
5	17.8	71.2	1800	3 AAC41738	Aac41738 Arabidops
6	17.6	70.4	1362	4 AAF60956	Aaf60956 P. putida
7	17.2	68.8	654	8 ACA21273	ACA21273 Prokaryot
8	17.2	68.8	5588	6 ABK63658	Abk63658 Rat seque
9	17.2	68.8	66499	13 ABD33094	Abd33094 Murine ca
10	17	68.0	4359	8 ADA69739	Ada69739 Rice gene
11	17	68.0	54037	13 ABD33323	Abd33323 Murine ca
12	17	68.0	59515	11 ACN44660	Acn44660 Mouse gen
13	17	68.0	73771	11 ACN44938	Acn44938 Human gen
14	17	68.0	126990	12 ADP13332	Adp13332 Renal cel
15	16.8	67.2	780	4 ABL16247	Abli16247 Drosophil
16	16.8	67.2	1248	8 ACA29348	Aca29348 Prokaryot
17	16.8	67.2	2838	4 ABL16246	Abli16246 Drosophil
18	16.8	67.2	3050	4 ABL18890	Abli18890 Drosophil
19	16.8	67.2	5956	4 ABL16242	Abli16242 Drosophil
20	16.8	67.2	110000	2 AAV30458_4	Continuation (5 of

21	16.8	67.2	110000	2 AAV30459_4	Continuation (5 of
22	16.6	66.4	45	2 AAQ5844	Aaq5844 Acetobact
23	16.6	66.4	45	2 AAQ86876	Aaq86876 Aat11R ge
24	16.6	66.4	349	10 ADE93875	Ade93875 Ixodes sc
25	16.6	66.4	447	4 AAH69833	Aah69833 Human cer
26	16.6	66.4	456	6 ABN60784	Abn60784 Human can
27	16.6	66.4	535	2 AAT83872	Aat83872 DNA encod
28	16.6	66.4	540	4 AAS57698	Aas57698 CDNA #374
29	16.6	66.4	1431	6 ABK35379	Abk35379 Human CDN
30	16.6	66.4	1492	8 ACA35711	Aca35711 Prokaryot
31	16.6	66.4	1518	4 AAS51408	Aas51408 Klebsiell
32	16.6	66.4	1518	8 ACA19360	Aca19360 Prokaryot
33	16.6	66.4	1521	4 AAS54023	Aas54023 Klebsiell
34	16.6	66.4	1539	11 ACH95348	Ach95348 Klebsiell
35	16.6	66.4	1812	2 AAT27105	Aat27105 Yeast cal
36	16.6	66.4	1826	12 ADH34512	Adh34512 CDNA enco
37	16.6	66.4	1849	2 AAZ20018	Aaz20018 Wheat 1,3
38	16.6	66.4	2053	12 ADH18989	Adh18989 Human cel
39	16.6	66.4	2085	4 AAS03081	Aas03081 Human dia
40	16.6	66.4	2149	10 ADD78284	Add78284 Human CGD
41	16.6	66.4	2204	8 ABX63768	Abx63768 Human CDN
42	16.6	66.4	3260	13 ADR07979	Adr07979 Full leng
43	16.6	66.4	4853	2 AAX01052	Aax01052 Nucleotid
44	16.6	66.4	35658	9 ADA02867	Ada02867 Mouse Sel
45	16.6	66.4	35658	10 ADB72605	Adb72605 Mouse Sel

ALIGNMENTS

RESULT 1  
ADJ39000/c

ID : ADJ39000 standard; DNA; 29751 BP.

XX

AC ADJ39000;

XX

DT 06-MAY-2004 (first entry)

XX

XX SARS coronavirus nucleotide sequence.

XX

small interfering RNA; siRNA; modified ribonucleotide;

XX viral replication inhibition; hepatitis C virus; HCV; hepatitis C;

XX antiinflammatory; hepatotropic; virucide; hepatitis A virus;

XX hepatitis D virus; hepatitis E virus; Ebola virus; influenza virus;

XX rotavirus; reovirus; retrovirus; poliovirus; human papilloma virus;

XX metapneumoniavirus; coronavirus; viral infection; gene; ds.

XX

OS SARS coronavirus.

XX

XX WO2004011647-A1.

XX

XX 05-FEB-2004.

XX

XX 25-JUL-2003; 2003WO-US023104.

XX

XX 26-JUL-2002; 2002US-0398605P.

XX

XX (CHIR ) CHIRON CORP.

XX

XX Han J, Seo MY, Houghton M;

XX

XX WPI; 2004-143862/14.

XX

PT New RNase resistant small interfering RNA, useful for treating viral

XX infections, e.g., hepatitis C, influenza virus or coronavirus infection.

XX

PS Example 10; Fig 3; 74pp; English.

XX

CC The present invention describes a small interfering RNA (siRNA) which

XX comprises a modified ribonucleotide, where the siRNA is resistant to

CC RNase and retains the ability to inhibit viral replication. Also

XX described: (1) inactivating a virus in a patient; (2) making a modified

CC siRNA that targets a nucleic acid sequence in a virus; (3) a double-

CC stranded RNA molecule of 10-30 nucleotides that inhibits replication of  
 CC hepatitis C virus (HCV); (4) inducing targeted RNA interference toward  
 CC HCV in hepatic cells; (5) inhibiting replication of HCV; (6) a vector  
 CC comprising a DNA segment encoding the RNA molecule; (7) a host cell  
 CC comprising the vector of (6); (8) inhibiting replication of HCV in cells  
 CC carrying HCV; (9) treating hepatitis C in a subject; (10) a modified  
 CC siRNA molecule comprising a double-stranded RNA molecule of 10-30  
 CC nucleotides in length, which mediates RNA interference toward a target  
 CC agent or virus and is linked to at least one receptor-binding ligand; and  
 CC (11) inducing targeted RNA interference in a patient. The modified siRNA  
 CC molecules have antiinflammatory, hepatotropic and virucide activities.  
 CC The modified RNA molecules are useful for inactivating virus in mammalian  
 CC cells. The siRNAs are useful for treating hepatitis C virus, hepatitis A  
 CC virus, hepatitis D virus, hepatitis E virus, Ebola virus, influenza  
 CC virus, rotavirus, reovirus, coronavirus, poliovirus, human papilloma  
 CC virus, metapneumoniavirus or coronavirus infections. The methods of the  
 CC invention can be used to correct or compensate for cellular physiological  
 CC abnormalities involved in conferring susceptibility to viral infections  
 CC in patients and/or alleviate symptoms of a viral infection in patients.  
 CC The present sequence represents the SARS coronavirus nucleotide sequence,  
 CC which is used in an example from the present invention.

XX  
 SQ Sequence 29751 BP; 8478 A; 5941 C; 6188 G; 9144 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 12; Length 29751;  
 Best Local Similarity 100.0%; Pred. No. 0.063; 0; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0;

QY 1 GCTGTGAACCAAGACGCGATTTAT 25  
 Db 28286 GCTGTGAACCAAGACGCGATTTAT 28262

## RESULT 2

ADN74528  
 ID ADN74528 standard; cDNA; 1533 BP.

XX  
 AC ADN74528;

XX  
 DT 15-JUL-2004 (first entry)

XX Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2423.

XX Gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;  
 KW growth regulator; animal feed product; thale cress;  
 KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX Arabidopsis thaliana.

XX WO2004035798-A2.

XX 29-APR-2004.

XX 20-OCT-2003; 2003WO-EP011658.

XX 18-OCT-2002; 2002EP-00079408.

XX (CROP-) CROPDESIGN NV.

XX Inze D, De Veylder L, Vlieghe K;

XX WPI; 2004-348466/32.

DR P-PSDB; ADN74529.

XX Altering plant characteristics, useful for producing plants for enzyme or  
 PT pharmaceutical production comprises modifying in a plant, expression of  
 PT one or more nucleic acids and/or modifying level or activity of one or  
 PT more proteins.

XX Claim 1; SEQ ID NO 2423; 134pp; English.

XX This invention relates to a novel method for altering one or more plant  
 CC characteristics. Specifically, it refers to identifying genes that are up

CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
 CC alter plant characteristics accordingly. The present invention describes  
 CC generating transgenic plants for the production of growth regulators,  
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
 CC the altered plant characteristics are selected from increased yield or  
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
 CC or physiology, altered endoreplication, biochemistry, signal  
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,  
 CC each relative to the corresponding wild type plants. Accordingly, these  
 CC sequences can also be useful as positive or negative selectable markers  
 CC during transformation of cells or tissues. The identified genes play a  
 CC role in a variety of biological processes such as DNA replication, cell  
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
 CC transcription factors. This polynucleotide sequence is thale cress cDNA  
 CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa  
 CC transcription factor, given in an exemplification of the invention.

XX  
 SQ Sequence 1533 BP; 478 A; 292 C; 352 G; 411 T; 0 U; 0 Other;

Query Match 71.2%; Score 17.8; DB 12; Length 1533;  
 Best Local Similarity 90.5%; Pred. No. 1.5e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 TGAACCAAGACGCGATTTAT 25

Db 1047 TGAACCAAGACGCGATTTAT 1067

## RESULT 3

AAC49448

ID AAC49448 standard; DNA; 1557 BP.

XX  
 AC AAC49448;

XX  
 DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 61198.

XX Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 16-APR-1999; 99US-0128455P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 23-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

XX 05-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 06-MAY-1999; 99US-0132487P.

XX 07-MAY-1999; 99US-0132863P.



PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135114P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136332P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140595P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-01445086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149426P.  
PR 18-AUG-1999; 99US-0149175P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149829P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 71.2%; Score 17.8; DB 3; Length 1557;  
Best Local Similarity 90.5%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 TGAACCAAGACGAGTATTAT 25  
||||||| ||||| ||||| ||||| |||||  
Db 1104 TGAACCAAGACGAGTATTAT 1124

RESULT 4  
AAC49462  
ID AAC49462 standard; DNA; 1792 BP.  
XX AAC49462;  
AC  
XX  
DT 18-OCT-2000 (first entry)  
XX  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 61250.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 08-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149303P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157717P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.

PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 71.2%; Score 17.8; DB 3; Length 1792;  
Best Local Similarity 90.5%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 TGAACCAAGACGAGTATTAT 25  
|||||  
Db 1122 TGAACCAAGACGAGTATTAT 1142  
  
RESULT 5  
AAC41738  
ID AAC41738 standard; DNA; 1800 BP.  
XX AAC41738;  
XX 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 32954.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 18-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 22-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161403P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match

71.2%; Score 17.8; DB 3; Length 1800;

Best Local Similarity 90.5%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 TGAACCAAGCAGCAGTATTAT 25  
||||| ||||| ||||| |||||  
Db 1120 TGAACCAAGCAGCAGTATTAT 1140

RESULT 6  
AAF60956  
ID AAF60956 standard; DNA; 1362 BP.  
XX  
AC AAF60956;  
XX  
DT 16-MAY-2001 (first entry)  
XX  
DE  
DE P. putida KT2440-associated DNA ORF00268.  
XX  
KW Transgenic plant; detection; probe; amplification; vaccine carrier;  
KW microbial production strain; biological remediation; ds.  
XX  
OS Pseudomonas putida.  
XX  
PN DE19935088-A1.  
XX  
PD 01-FEB-2001.  
XX  
XX 27-JUL-1999; 99DE-01035088.  
XX  
XX 27-JUL-1999; 99DE-01035088.  
XX  
XX (TIGR-) TIGR INST GENOMIC RES.  
PA (QUIA-) QUIAGEN GMBH.  
PA (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.  
PA (DKFZ-) DKFZ DEUT KREBSFORSCHUNGSZENTRUM.  
PA (MEDI-) MEDIZINISCHE HOCHSCHULE HANNOVER.  
XX  
XX WPI; 2001-192469/20.  
XX  
XX New DNA sequences specific for Pseudomonas putida KT2440, useful as safe  
PT genetic engineering host, allow detection in presence of other related  
PT bacteria.  
XX  
PS Claim 1a; Page 4; 158pp; German.  
XX  
CC This invention describes novel DNA sequences (I) for specific detection  
CC of Pseudomonas putida KT2440. The invention also describes (1)  
CC recombinant expression vector containing (I); (2) prokaryotic or  
CC eukaryotic cells transformed or transfected with (I) or the vector of (1)  
CC; (3) production of expression products by culturing cells of (2); (4)  
CC expression products, or their fragments, of (1) and synthetic proteins or  
CC peptides with the same sequences (A); (5) poly- or mono-clonal antibodies  
CC (Ab) that react specifically with (A); (6) hybridoma cells that produce  
CC the monoclonal Ab of (5); (7) transgenic plants that contain transformed  
CC or transfected cells of (2); (8) detecting KT2440 using a labeled (I) or  
CC Ab as probe; and (9) DNA chips carrying one or more (I). (I), and their  
CC fragments, are used as probes to detect and isolate full-length cDNAs  
CC and/or to amplify such cDNAs by polymerase chain reaction, and for  
CC production of transgenic plants. (I), or antibodies that recognize their  
CC expression products, are used for detecting the presence of KT2440, or  
CC particularly in presence of other, even closely related, bacteria. KT2440  
CC is one of the bacteria classified as safe, by the National Institutes of  
CC Health, for genetic engineering work, e.g. as microbial production  
CC strains, for biological remediation and as vaccine carriers. (I) are  
CC exclusive to KT2440 with no significant homology with sequences in other  
CC bacteria (specifically the closely related pathogen P. aeruginosa).  
CC Compared with other 'safe' bacteria, it has greater catabolic activity  
CC and better survival in, and adaptation to, the rhizosphere and soil  
XX  
SQ Sequence 1362 BP; 316 A; 376 C; 382 G; 288 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 4; Length 1362;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CTGTGAACCAAGCAGCAGTATTAT 25  
||||| ||||| ||||| |||||  
Db 1237 CTGTGAACCAAGCAGCAGTATTAT 1260

RESULT 7  
ACA21273  
ID ACA21273 standard; DNA; 654 BP.  
XX  
AC ACA21273;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Prokaryotic essential gene #2930.  
XX  
KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
XX  
OS Bacillus anthracis.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
XX  
XX 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
PI  
XX WPI; 2003-029926/02.  
DR P-PSDB; ABU17403.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 9143; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 654 BP; 240 A; 89 C; 147 G; 178 T; 0 U; 0 Other;  
Query Match 68.8%; Score 17.2; DB 8; Length 654;  
Best Local Similarity 86.4%; Pred. No. 2.6e+02;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 GTGAACCAAGACGCGAGTATTAT 25  
||||| ||||||| |||||||  
Db 476 GTGAATCAAGACGCGATTATTTT 497

RESULT 8  
ABK63658  
ID ABK63658 standard; cDNA; 5588 BP.

AC ABK63658;

DT 18-JUN-2002 (first entry)

XX Rat sequence differentially expressed in response to a hepatotoxin #1565.  
DE  
XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;  
KW differential expression; centrilobular necrosis; steatosis.  
XX

OS *Rattus norvegicus*.

XX WO200210453-A2.

PN 07-FEB-2002.

PD

XX 30-JUL-2001; 2001WO-US023872.

PR 31-JUL-2000; 2000US-0222040P.

PR 02-NOV-2000; 2000US-0244880P.

PR 11-MAY-2001; 2001US-0290029P.

PR 15-MAY-2001; 2001US-0290645P.

PR 22-MAY-2001; 2001US-0292336P.

PR 06-JUN-2001; 2001US-0295798P.

PR 13-JUN-2001; 2001US-0297457P.

PR 19-JUN-2001; 2001US-0298884P.

PR 09-JUL-2001; 2001US-0303459P.

XX (GENE-) GENE LOGIC INC.

PA Mendrick D, Porter MW, Johnson KR, Castle AL, Blaschhoff MR;

XX WPI; 2002-241625/29.

XX Predicting toxic effects of compounds or the progression of these toxic  
PT effects by determining the changes in gene expression in tissues or cells  
PT exposed to the toxin and comparing these to gene expression in unexposed  
PT tissues or cells.  
XX Claim 1; SEQ ID NO 1565; 239pp; English.  
PS The invention relates to methods for predicting toxic effects of  
CC compounds or the progression of these toxic effects by determining the  
CC global changes in gene expression in tissues or cells exposed to the  
CC toxin and comparing these to gene expression in unexposed tissues or  
CC cells. Also included are methods of predicting at least one toxic effect  
CC of a compound or progression of a toxic effect, preferably the  
CC hepatotoxicity of a compound, comprising detecting the level of  
CC expression in a tissue or cell sample exposed to the compound of two or  
CC more genes listed in the specification, where differential expression of  
CC the genes is indicative of at least one toxic effect or progression. The  
CC method can also be used to identify an agent which modulates the toxic

CC response and predict cellular pathways that a compound modulates in a  
CC cell. The methods utilise a set of at least two probes (on a solid  
CC support in kit form), where each of the probes comprises a sequence that  
CC specifically hybridises to a gene listed in the specification, a computer  
CC system comprising a database containing information identifying the  
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a  
CC set of genes comprising at least two genes listed in the specification,  
CC and a user interface to view the information used to present information  
CC identifying the expression level in a tissue or cell of at least one gene  
CC listed in the specification. The method is useful for elucidating global  
CC changes in gene expression and for identifying toxicity markers in  
CC tissues or cell exposed to a known toxin. The genes may be used as  
CC toxicity markers in drug screening and toxicity assays. The genes and  
CC gene expression information may be used as diagnostic markers for the  
CC prediction or identification of the physiological state of tissue or cell  
CC sample that has been exposed to a compound or agent. Hepatotoxicity is  
CC characterised by centrilobular necrosis and steatosis. The present  
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
CC which is differentially expressed in response to a hepatotoxic agent

XX Sequence 5588 BP; 1349 A; 1550 C; 1416 G; 1272 T; 0 U; 1 Other;

Query Match 68.8%; Score 17.2; DB 6; Length 5588;

Best Local Similarity 86.4%; Pred. No. 3.6e+02;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CTGTGAACCAAGACGCGAGTATT 23

||||| ||||||| |||||||  
Db 1216 CTGAGAACCAAGACGCGAGAGT 1237

RESULT 9

ABD33094

ID ABD33094 standard; DNA; 66499 BP.

XX ABD33094;

XX 18-NOV-2004 (first entry)

DE Murine cancer-associated (CA) gene MD07-007.

XX Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;

XX ds; cancer; cytostatic.

XX Mus musculus.

XX WO2004058146-A2.

XX 15-JUL-2004.

XX 15-DEC-2003; 2003WO-US040081.

XX 17-DEC-2002; 2002US-00322281.

XX (SAGR-) SAGRES DISCOVERY INC.

XX Morris DW, Malandro MS;

XX WPI; 2004-499109/47.

PT Novel human cancer associated protein encoded within open reading frame  
PT of cancer associated gene, useful as targets for diagnosing cancer.

XX Disclosure; SEQ ID NO 39; 182pp; English.

XX The invention relates to cancer-associated proteins (CAP) and the cancer-  
XX associated (CA) nucleic acids encoding them. The invention also relates  
XX to a method for treating cancers involving administering to a patient an  
XX inhibitor of CAP, and a method of screening for anticancer activity in a  
XX potential drug involving providing a cell that expresses a CA gene,  
XX contacting a tissue sample derived from a cancer cell with an anticancer  
XX drug candidate and monitoring the effect of the anticancer drug candidate  
XX on expression of the CA gene. The CAP proteins are useful for detecting

CC cancer associated with expression of a CAP protein in a test cell sample  
CC and for screening for a bioactive agent capable of modulating the  
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing  
CC cancer, involving determining the expression of a CA nucleic acid in a  
CC tissue. This sequence represents a murine CA gene of the invention. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 66499 BP; 16137 A; 16098 C; 16364 G; 17399 T; 0 U; 501 Other;  
Query Match 68.8%; Score 17.2; DB 13; Length 66499;  
Best Local Similarity 86.4%; Pred. No. 5.3e+02;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CTGTGAACCAAGACGAGTATT 23  
||||| ||||| ||||| ||||| |||||  
Db 4519 CTGTGACCAAGACACAGATT 4540

RESULT 10  
ADA69739  
ID ADA69739 standard; DNA; 4359 BP.

XX AC ADA69739;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 3062.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175290/17.

XX PT Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.

XX PS Claim 6; SEQ ID NO 3062; 899pp; English.

XX CC The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.

XX SQ Sequence 4359 BP; 1100 A; 856 C; 1130 G; 1271 T; 0 U; 2 Other;

Query Match 68.0%; Score 17; DB 8; Length 4359;

Best Local Similarity 80.0%; Pred. No. 4.4e+02;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGAGTATT 25  
||||| ||||| ||||| ||||| |||||  
Db 3354 GATGGGAACGAACACATTATT 3378

RESULT 11

ABD33323

ID ABD33323 standard; DNA; 54037 BP.

XX AC ABD33323;

XX DT 18-NOV-2004 (first entry)

XX DE Murine cancer-associated (CA) gene MD07-058.

XX KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;  
XX ds; cancer; cytostatic.

XX OS Mus musculus.

XX PN WO2004058146-A2.

XX PD 15-JUL-2004.

XX PF 15-DEC-2003; 2003WO-US040081.

XX PR 17-DEC-2002; 2002US-00322281.

XX PA (SAGR-) SAGRES DISCOVERY INC.

XX PI Morris DW, Malandro MS;

XX DR WPI; 2004-499109/47.

XX PT Novel human cancer associated protein encoded within open reading frame  
XX of cancer associated gene, useful as targets for diagnosing cancer.

XX PS Disclosure; SEQ ID NO 395; 182pp; English.

XX CC The invention relates to cancer-associated proteins (CAP) and the cancer-  
XX associated (CA) nucleic acids encoding them. The invention also relates  
XX to a method for treating cancers involving administering to a patient an  
XX inhibitor of CAP, and a method of screening for anticancer activity in a  
XX potential drug involving providing a cell that expresses a CA gene,  
XX contacting a tissue sample derived from a cancer cell with an anticancer  
XX drug candidate and monitoring the effect of the anticancer drug candidate  
XX on expression of the CA gene. The CAP proteins are useful for detecting  
XX cancer associated with expression of a CAP protein in a test cell sample  
XX and for screening for a bioactive agent capable of modulating the  
XX activity of a CAP protein. The CA nucleic acids are useful for diagnosing  
XX cancer, involving determining the expression of a CA nucleic acid in a  
XX tissue. This sequence represents a murine CA gene of the invention. Note:  
XX The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 54037 BP; 13435 A; 12064 C; 12935 G; 15319 T; 0 U; 284 Other;

Query Match 68.0%; Score 17; DB 13; Length 54037;

Best Local Similarity 80.0%; Pred. No. 6.4e+02;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGAGTATT 25

Db 35534 GCAGTGAACCATGACGAAGTGTGT 35558

RESULT 12

ACN44660/c

ID ACN44660 standard; DNA; 69515 BP.

XX AC ACN44660;

```
XX 18-NOV-2004 (first entry)
DT Mouse genomic sequence mCG22656.
DE Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX Mus musculus.
XX WO2003073826-A2.
XX 12-SEP-2003.
XX 28-FEB-2003; 2003WO-US006235.
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
XX Claim 1; SEQ ID NO 1219; Opp; English.
XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for treating
XX carcinoma; (vi) for inhibiting the activity of CAP; (ix) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published
XX
XX Sequence 69515 BP; 17960 A; 14338 C; 15207 G; 21693 T; 0 U; 317 Other;
XX
XX Query Match 68.0%; Score 17; DB 11; Length 69515;
XX Best Local Similarity 80.0%; Pred. No. 6.7e+02;
XX Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 GCTGTGACCAAGACGCGAGTATTAT 25
XX 21903 GCTGTGAAACAGACATTGTATAT 21879
XX
XX RESULT 13
XX ACN44938
XX ID ACN44938 standard; DNA; 73771 BP.
XX AC ACN44938;
XX 18-NOV-2004 (first entry)
XX Human genomic sequence hCG37394.
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX Homo sapiens.
XX WO2003073826-A2.
XX 12-SEP-2003.
XX 28-FEB-2003; 2003WO-US006235.
XX
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
XX Claim 1; SEQ ID NO 1219; Opp; English.
XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for treating
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published
XX
XX Sequence 73771 BP; 20930 A; 14938 C; 16429 G; 21180 T; 0 U; 294 Other;
XX
XX Query Match 68.0%; Score 17; DB 11; Length 73771;
XX Best Local Similarity 80.0%; Pred. No. 6.7e+02;
XX Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 GCTGTGACCAAGACGCGAGTATTAT 25
XX 49607 GCTGTGATTCAGAGGCGAGATTAT 49631
XX
XX RESULT 14
XX ADP13332
XX ID ADP13332 standard; DNA; 126990 BP.
XX AC ADP13332;
XX 26-AUG-2004 (first entry)
XX Renal cell carcinoma differentially expressed gene #68.
XX ds; diagnosis; non-blood disease; solid tumor; gene expression;
XX peripheral blood mononuclear cell; renal cell carcinoma; prostate cancer;
XX head/neck cancer; differential expression.
XX Homo sapiens.
XX WO2004048933-A2.
XX 10-JUN-2004.
XX 21-NOV-2003; 2003WO-US037481.
XX 21-NOV-2002; 2002US-0427982P.
XX 03-APR-2003; 2003US-0459782P.
XX (AMHP ) WVETH.
XX (TWIN/) TWINE N C.
XX (BURC/) BURCZYNSKI M E.
XX (TREP/) TREPICCHIO W L.
XX (DORN/) DORNER A.
XX (STOV/) STOVER J A.
XX (SLON/) SLONI D K.
XX
```



PI Twine NC, Burczynski ME, Trepicchio WL, Dorner A, Stover JA;  
PI Stoni DK;  
XX WPI; 2004-460799/43.  
XX  
XX Diagnosing non-blood disease such as solid tumor, involves comparing  
PT differential expression profile of specific genes in peripheral blood  
PT sample of subject with reference expression profile of specific genes.  
XX  
XX Disclosure; SEQ ID NO 68; 350pp; English.  
XX  
XX The invention relate to a method of diagnosing (M1) non-blood disease  
CC such as solid tumor by providing peripheral blood sample of human having  
CC non-blood disease, and comparing an expression profile of specific genes  
CC in the peripheral blood sample to reference expression profile of the  
CC genes, where each of the genes is differentially expressed in peripheral  
CC blood mononuclear cells (PBMCs) of patients having the disease as  
CC compared to PBMCs of normal humans. The method is useful for diagnosing  
CC non-blood disease such as solid tumor. The solid tumor is chosen from  
CC renal cell carcinoma (RCC), prostate cancer and head/neck cancer. The  
CC peripheral blood sample comprises enriched PBMCs. The peripheral blood  
CC sample is a whole blood sample (claimed). (M1) is useful for identifying  
CC genes that are differentially expressed in peripheral blood samples  
CC isolated at different stages of progression, development or treatment of  
CC RCC and/or other solid tumors. This sequence corresponds to a gene that  
CC is differentially expressed and detected by the method of the invention.  
CC (Note: this sequence is not given as part of the printed specification  
CC but was obtained from WIPO in electronic format at  
CC ftp.wipo./pub/published\_pct\_sequences).  
XX  
XX Sequence 126990 BP; 36683 A; 27389 C; 27065 G; 35853 T; 0 U; 0 Other;  
SQ  
Query Match 68.0%; Score 17; DB 12; Length 126990;  
Best Local Similarity 80.0%; Pred. No. 7.3e+02;  
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 GCTGTGACCAAGCAGCAGTATTAT 25  
DB 108520 GCTGTGACCAAGCAGCAGTATTAT 108544  
RESULT 15  
ABL16247/c  
ID ABL16247 standard; DNA; 780 BP.  
XX  
XX AC ABL16247;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 214.  
DE  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
OS  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR  
XX 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.  
XX  
XX Claim 1; SEQ ID NO 214; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 780 BP; 206 A; 221 C; 199 G; 154 T; 0 U; 0 Other;  
SQ  
Query Match 67.2%; Score 16.8; DB 4; Length 780;  
Best Local Similarity 90.0%; Pred. No. 4.2e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GCTGTGACCAAGCAGCAGT 20  
DB 201 GCTGTGACCAAGCAGCAGT 182  
Search completed: May 16, 2005, 01:09:35  
Job time : 324.703 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:11:05 ; Search time 93.5547 Seconds  
(without alignments)  
437.251 Million cell updates/sec

Title: US-10-808-187A-2475

Perfect score: 25

Sequence: 1 gctgtgaacacgacgagctattat 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.6	70.4	46492	4	US-09-949-016-12953
2	17.6	70.4	46492	4	US-09-949-016-12954
3	17	68.0	357	4	US-09-248-796A-8219
4	17	68.0	51336	4	US-09-949-016-16054
5	17	68.0	109025	4	US-09-949-016-12609
6	17	68.0	109025	4	US-09-949-016-17567
7	16.8	67.2	107800	4	US-09-949-016-13118
8	16.8	67.2	116425	4	US-09-949-016-11809
9	16.8	67.2	536165	4	US-09-214-808-1
10	16.6	66.4	45	1	US-08-134-570-3
11	16.6	66.4	1455	4	US-09-248-796A-6254
12	16.6	66.4	1539	4	US-09-489-039A-1143
13	16.6	66.4	1812	1	US-08-328-322-18
14	16.6	66.4	1964	1	US-08-328-322-9
15	16.6	66.4	2034	4	US-09-949-016-2614
16	16.6	66.4	2034	4	US-09-949-016-2939
17	16.6	66.4	4853	3	US-08-881-450A-22
18	16.6	66.4	38298	4	US-09-949-016-16650
19	16.6	66.4	95750	4	US-09-949-016-11926
20	16.6	66.4	202001	4	US-09-734-674-3
21	16.6	66.4	218940	4	US-09-949-016-17539
22	16.2	64.8	87	3	US-08-976-413A-335
23	16.2	64.8	15084	4	US-09-949-016-16277
24	16.2	64.8	43981	4	US-09-949-016-16294
25	16	64.0	601	4	US-09-949-016-59665
26	16	64.0	1677	4	US-09-328-352-744
27	16	64.0	1875	4	US-09-543-681A-171

28	16	64.0	2763	4	US-09-463-402-5	Sequence 5, Appli
29	16	64.0	2766	4	US-09-889-572-3	Sequence 3, Appli
30	16	64.0	3404	4	US-09-710-279-4346	Sequence 4346, Ap
31	16	64.0	4065	4	US-09-463-402-13	Sequence 13, Appli
32	16	64.0	5860	4	US-09-949-016-4463	Sequence 4463, Ap
33	16	64.0	8279	4	US-09-949-016-13478	Sequence 13478, A
34	16	64.0	9531	4	US-09-163-748C-3	Sequence 3, Appli
35	16	64.0	11461	3	US-08-669-161A-29	Sequence 29, Appli
36	16	64.0	15473	4	US-09-949-016-12851	Sequence 12851, A
37	16	64.0	15474	4	US-09-949-016-13890	Sequence 13890, A
38	16	64.0	24979	2	US-08-147-777-3	Sequence 3, Appli
39	16	64.0	24979	3	US-08-452-872-3	Sequence 3, Appli
40	16	64.0	24979	5	PCT-US93-03985-3	Sequence 3, Appli
41	16	64.0	32495	4	US-09-949-016-14952	Sequence 14952, A
42	16	64.0	51629	4	US-09-949-016-12883	Sequence 12883, A
43	16	64.0	51631	4	US-09-949-016-15496	Sequence 15496, A
44	16	64.0	57875	4	US-09-949-016-13152	Sequence 13152, A
45	16	64.0	103934	4	US-09-949-016-14433	Sequence 14433, A

ALIGNMENTS

RESULT 1  
US-09-949-016-12953  
; Sequence 12953, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12953  
; LENGTH: 46492  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(46492)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12953  
Query Match 70.4%; Score 17.6; DB 4; Length 46492;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 GCTGTGAACCAACGACGAGTATTA 24  
|||||  
Db 7003 GCTGGGACCCAGCGAGTATCA 7026  
RESULT 2  
US-09-949-016-12954  
; Sequence 12954, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12954  
; LENGTH: 4692  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(46492)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12954

Query Match 70.4%; Score 17.6; DB 4; Length 46492;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTATTA 24  
||||| ||||| ||||| ||||| |||||  
Db 7003 GCTGGGAACCCAGGCGCGAGTATCA 7026

## RESULT 3

US-09-248-796A-8219  
; Sequence 8219, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 8219  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (36)  
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknc

Query Match 68.0%; Score 17; DB 4; Length 357;  
Best Local Similarity 80.0%; Pred. No. 97;  
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25  
||||| ||||| ||||| ||||| |||||  
Db 232 GATGTGCACCAAGACGCGAGTAAAGAT 256

## RESULT 4

US-09-949-016-16054  
; Sequence 16054, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16054  
; LENGTH: 5136  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(5136)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16054

Query Match 68.0%; Score 17; DB 4; Length 51336;  
Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25  
||||| ||||| ||||| ||||| |||||  
Db 35840 GCTGTGATTCAGGAGGCGAGATTAT 35864

## RESULT 5

US-09-949-016-12609  
; Sequence 12609, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12609  
; LENGTH: 109025  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(109025)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12609

Query Match 68.0%; Score 17; DB 4; Length 109025;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25  
||||| ||||| ||||| ||||| |||||  
Db 32583 GTTGTGAACCTGAGACCCAGGATTAT 32607

## RESULT 6

US-09-949-016-17567  
; Sequence 17567, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17567  
; LENGTH: 109025  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(109025)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17567

Query Match 68.0%; Score 17; DB 4; Length 109025;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCAGCAGTATTAT 25  
DB 32583 GTTGTGAAGTGAAGCAGCAGGATTAT 32607

RESULT 7  
US-09-949-016-13118/c  
; Sequence 13118, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13118  
; LENGTH: 107800  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13118  
Query Match 67.2%; Score 16.8; DB 4; Length 107800;  
Best Local Similarity 90.0%; Pred. No. 3.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GCTGTGAACCAAGCAGCAGT 20  
DB 20152 GCTGTGAACCAAGAGCCGT 20133  
RESULT 8  
US-09-949-016-11809/c  
; Sequence 11809, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11809  
; LENGTH: 116425  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11809

Query Match 67.2%; Score 16.8; DB 4; Length 116425;  
Best Local Similarity 90.0%; Pred. No. 3.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCAGCAGT 20  
DB 28778 GCTGTGAACCAAGAGCCGT 28759

RESULT 9  
US-09-214-808-1  
; Sequence 1, Application US/09214808A  
; Patent No. 6475793  
; GENERAL INFORMATION:  
; APPLICANT: Rosenthal, Andre  
; APPLICANT: Freiberg, Christophe  
; APPLICANT: Perret, Xavier Philippe  
; APPLICANT: Broughton, William John  
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic  
; Patent No. 6475793  
; TITLE OF INVENTION: Plasmid  
; FILE REFERENCE: CARP0088  
; CURRENT APPLICATION NUMBER: US/09/214,808A  
; CURRENT FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: PCT/IB97/00950  
; PRIOR FILING DATE: 1997-07-10  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 536165  
; TYPE: DNA  
; ORGANISM: Rhizobium  
US-09-214-808-1  
Query Match 67.2%; Score 16.8; DB 4; Length 536165;  
Best Local Similarity 90.0%; Pred. No. 4.3e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GTGAACCAAGCAGCAGTATT 23  
DB 416185 GTGACCAAGAGCAGTATT 416204

RESULT 10  
US-08-134-570-3  
; Sequence 3, Application US/08134570  
; Patent No. 5405768  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Shuang-yong  
; APPLICANT: Nwankwo, Donald O.  
; APPLICANT: Forney, Julie A.  
; APPLICANT: Wilson, Geoffrey G.  
; APPLICANT: Zhang, Bohong  
; TITLE OF INVENTION: METHOD FOR CLONING AND PRODUCING THE  
; AATII AND ALII RESTRICTION ENDONUCLEASE AND METHYLASE  
; TITLE OF INVENTION: AND RELATED METHOD FOR OVEREXPRESSION RESTRICTION  
; TITLE OF INVENTION: ENDONUCLEASES  
; NUMBER OF SEQUENCES: 16

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &  
;; ADDRESSEE: CUSHMAN  
;; STREET: 130 WATER STREET  
;; CITY: BOSTON  
;; STATE: MASSACHUSETTS  
;; COUNTRY: US  
;; ZIP: 02109-4280  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/134,570  
;; FILING DATE: 10-08-93  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/909947  
;; FILING DATE: 07-JUL-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: CORLESS, PETER F.  
;; REGISTRATION NUMBER: 33,860  
;; REFERENCE/DOCKET NUMBER: 41,891-CIP  
;; TELEPHONE: (617) 523-3400  
;; TELEFAX: (617) 523-6440  
;; TELEX: 200291 STRE UR  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 45 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
US-08-134-570-3

Query Match 66.4%; Score 16.6; DB 1; Length 45;  
Best Local Similarity 82.6%; Pred. No. 1e+02;  
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGTGAACCAAGCGCAGTATT 25  
DB 21 TATGAACCCAGACGAAGTATTT 43

RESULT 11  
US-09-248-796A-6254/c  
; Sequence 6254, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 6254  
; LENGTH: 1455  
; TYPE: DNA  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (21)  
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknd  
US-09-248-796A-6254

Query Match 66.4%; Score 16.6; DB 4; Length 1455;  
Best Local Similarity 82.6%; Pred. No. 2e+02;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 2 CTGTGAACCAAGCGCAGTATTA 24  
DB 73 CTGTGAACCAAGACAGATATATTA 51

RESULT 12  
US-09-489-039A-1143/c  
; Sequence 1143, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 1143  
; LENGTH: 1539  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-1143

Query Match 66.4%; Score 16.6; DB 4; Length 1539;  
Best Local Similarity 82.6%; Pred. No. 2e+02;  
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCGCAGTATT 23  
DB 1531 GCAATGACCAAGACGCGAGTATT 1509

RESULT 13  
US-08-328-322-18  
; Sequence 18, Application US/08328322  
; Patent No. 5723436  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Laiqiang  
; APPLICANT: Cyert, Martha S.  
; TITLE OF INVENTION: Calcineurin Interacting Protein Compositions  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/328,322  
; FILING DATE: 24-OCT-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: P38,615  
; REFERENCE/DOCKET NUMBER: 8600-0151.10  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1812 base pairs  
; TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: DNA fragment containing  
INDIVIDUAL ISOLATE: CNAIdeltaC coding sequence  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 286..1812  
US-08-328-322-18

Query Match 66.4%; Score 16.6; DB 1; Length 1812;  
Best Local Similarity 82.6%; Pred. No. 2.1e+02;  
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGTGAACCAAGCGCAGTATTAT 25  
DB 141 TTTGAATCAAGAGCGCATTTAT 163

RESULT 14  
US-08-328-322-9  
Sequence 9, Application US/08328322  
Patent No. 5723436  
GENERAL INFORMATION:  
APPLICANT: Huang, Laiqiang  
APPLICANT: Cyert, Martha S.  
TITLE OF INVENTION: Calcineurin Interacting Protein Compositions  
TITLE OF INVENTION: and Methods  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/328,322  
FILING DATE: 24-OCT-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: P38,615  
REFERENCE/DOCKET NUMBER: 8600-0151.10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1964 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: genomic DNA fragment containing full  
INDIVIDUAL ISOLATE: CNA1 coding sequence  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 286..1944  
US-08-328-322-9

Query Match 66.4%; Score 16.6; DB 1; Length 1964;

Best Local Similarity 82.6%; Pred. No. 2.1e+02;  
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 3 TGTGAACCAAGCGCAGTATTAT 25  
DB 141 TTTGAATCAAGAGCGCATTTAT 163

RESULT 15  
US-09-949-016-2614/c  
Sequence 2614, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2614  
LENGTH: 2034  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-2614

Query Match 66.4%; Score 16.6; DB 4; Length 2034;  
Best Local Similarity 82.6%; Pred. No. 2.1e+02;  
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CTGTGAACCAAGCGCAGTATTA 24  
DB 934 CTGCAAAACCAAGACACAGGATTA 912

Search completed: May 16, 2005, 06:11:44  
Job time : 97.5547 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 02:00:36 ; Search time 1185.16 Seconds  
(without alignments)  
129.102 Million cell updates/sec

Title: US-10-808-187A-2475

Perfect score: 25

Sequence: 1 gctgtgacacagcagcagtattat 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	19	US-10-808-187-2475
2	25	100.0	1620	19	US-10-699-936-16
3	25	100.0	2304	19	US-10-699-936-7
4	25	100.0	2304	19	US-10-699-936-11
5	25	100.0	24774	19	US-10-889-447-3
6	25	100.0	28920	19	US-10-889-447-5
7	25	100.0	28920	19	US-10-889-447-6
8	25	100.0	29291	19	US-10-889-447-4
9	25	100.0	29430	19	US-10-889-447-7
10	25	100.0	29727	18	US-10-839-729-15
11	25	100.0	29727	18	US-10-827-757-1

c 12	25	100.0	29727	19	US-10-889-447-8	Sequence 8, Appli
c 13	25	100.0	29727	19	US-10-699-936-1	Sequence 1, Appli
c 14	25	100.0	29736	18	US-10-839-729-17	Sequence 17, Appli
c 15	25	100.0	29736	19	US-10-889-447-9	Sequence 9, Appli
c 16	25	100.0	29736	19	US-10-699-936-3	Sequence 3, Appli
c 17	25	100.0	29742	18	US-10-839-729-16	Sequence 16, Appli
c 18	25	100.0	29742	19	US-10-808-187-15	Sequence 15, Appli
c 19	25	100.0	29742	19	US-10-808-187-16	Sequence 16, Appli
c 20	25	100.0	29742	19	US-10-808-187-240	Sequence 240, App
c 21	25	100.0	29742	19	US-10-808-187-737	Sequence 737, App
c 22	25	100.0	29742	19	US-10-808-187-1108	Sequence 1108, Ap
c 23	25	100.0	29742	19	US-10-808-187-1590	Sequence 1590, Ap
c 24	25	100.0	29742	19	US-10-808-187-1365	Sequence 1365, Ap
c 25	25	100.0	29742	19	US-10-889-447-10	Sequence 10, Appli
c 26	25	100.0	29751	18	US-10-839-729-14	Sequence 14, Appli
c 27	25	100.0	29751	19	US-10-856-529-1	Sequence 1, Appli
c 28	25	100.0	29751	19	US-10-626-879-67	Sequence 67, Appli
c 29	25	100.0	29751	19	US-10-889-447-1	Sequence 1, Appli
c 30	25	100.0	29751	19	US-10-889-447-2	Sequence 2, Appli
c 31	25	100.0	29751	19	US-10-699-936-2	Sequence 2, Appli
c 32	18.8	75.2	270	18	US-10-425-115-116691	Sequence 116691,
c 33	18.4	73.6	1386	17	US-10-425-114-6765	Sequence 6765, Ap
c 34	18.2	72.8	2278	13	US-10-027-632-110360	Sequence 110360,
c 35	18.2	72.8	2278	17	US-10-027-632-110360	Sequence 110360,
c 36	17.8	71.2	1930	17	US-10-425-114-29340	Sequence 29340, A
c 37	17.8	71.2	1972	17	US-10-425-114-14784	Sequence 14784, A
c 38	17.8	71.2	1972	17	US-10-424-599-46514	Sequence 46514, A
c 39	17.6	70.4	117	17	US-10-424-599-115998	Sequence 115998,
c 40	17.6	70.4	1097	18	US-10-767-701-3584	Sequence 3584, Ap
c 41	17.2	68.8	593	13	US-10-027-632-188195	Sequence 188195,
c 42	17.2	68.8	593	17	US-10-027-632-188195	Sequence 188195,
c 43	17.2	68.8	654	17	US-10-282-122A-9143	Sequence 9143, Ap
c 44	17.2	68.8	5588	9	US-09-917-800A-1565	Sequence 1565, Ap
c 45	17.2	68.8	66499	18	US-10-322-281-39	Sequence 39, Appli

#### ALIGNMENTS

#### RESULT 1

US-10-808-187-2475  
; Sequence 2475, Application US/10808187  
; Publication No. US2005000909A1  
; GENERAL INFORMATION:  
; APPLICANT: PEIRIS, JOSEPH S. M.  
; APPLICANT: YUEN, KWOK YUNG  
; APPLICANT: POON, LIT MAN  
; APPLICANT: GUAN, YI  
; APPLICANT: CHAN, KWOK HUNG  
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE  
; TITLE OF INVENTION: RESPIRATORY SYNDROME (SARS)  
; FILE REFERENCE: V9661.0078  
; CURRENT APPLICATION NUMBER: US/10/808,187  
; CURRENT FILING DATE: 2004-03-24  
; PRIOR APPLICATION NUMBER: 60/457,031  
; PRIOR FILING DATE: 2003-03-24  
; PRIOR APPLICATION NUMBER: 60/457,730  
; PRIOR FILING DATE: 2003-03-26  
; PRIOR APPLICATION NUMBER: 60/459,931  
; PRIOR FILING DATE: 2003-04-02  
; PRIOR APPLICATION NUMBER: 60/460,357  
; PRIOR FILING DATE: 2003-04-03  
; PRIOR APPLICATION NUMBER: 60/461,265  
; PRIOR FILING DATE: 2003-04-08  
; PRIOR APPLICATION NUMBER: 60/462,805  
; PRIOR FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 60/468,139  
; PRIOR FILING DATE: 2003-05-05  
; PRIOR APPLICATION NUMBER: 60/464,886  
; PRIOR FILING DATE: 2003-04-23  
; PRIOR APPLICATION NUMBER: 60/471,200  
; PRIOR FILING DATE: 2003-05-16

```
; NUMBER OF SEQ ID NOS: 2476
; SOFTWARE: PatentIn ver. 3.2
; SEQ ID NO 2475
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-808-187-2475
```

```
Query Match      100.0%; Score 25; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCTGTGAACCAAGACGAGCGAGTATTAT 25
    |||||||
DB 1 GCTGTGAACCAAGACGAGCGAGTATTAT 25
```

```
RESULT 2
US-10-699-936-16/c
; Sequence 16, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: SARS-CoV ZJ-HZ01
US-10-699-936-16
```

```
Query Match      100.0%; Score 25; DB 19; Length 1620;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCTGTGAACCAAGACGAGCGAGTATTAT 25
    |||||||
DB 241 GCTGTGAACCAAGACGAGCGAGTATTAT 217
```

```
RESULT 3
US-10-699-936-7/c
; Sequence 7, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 2304
; TYPE: DNA
; ORGANISM: SARS-CoV Shanghai LY
US-10-699-936-7
```

```
Query Match      100.0%; Score 25; DB 19; Length 2304;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCTGTGAACCAAGACGAGCGAGTATTAT 25
    |||||||
DB 921 GCTGTGAACCAAGACGAGCGAGTATTAT 897
```

```
RESULT 4
US-10-699-936-11/c
; Sequence 11, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 2304
; TYPE: DNA
; ORGANISM: SARS coronavirus Shanghai LY
US-10-699-936-11
```

```
Query Match      100.0%; Score 25; DB 19; Length 2304;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCTGTGAACCAAGACGAGCGAGTATTAT 25
    |||||||
DB 921 GCTGTGAACCAAGACGAGCGAGTATTAT 897
```

```
RESULT 5
US-10-889-447-3/c
; Sequence 3, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 24774
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-3
```

```
Query Match      100.0%; Score 25; DB 19; Length 24774;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCTGTGAACCAAGACGAGCGAGTATTAT 25
    |||||||
```

```
Db 23399 GCTGTGAACCAAGACGCGAGTATTAT 23375
RESULT 6
US-10-889-447-5/c
; Sequence 5, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; PRIOR FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ03
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-5
Query Match 100.0%; Score 25; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25
|||||
Db 27548 GCTGTGAACCAAGACGCGAGTATTAT 27524

RESULT 7
US-10-889-447-6/c
; Sequence 6, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; PRIOR FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ04
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-6
Query Match 100.0%; Score 25; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25
|||||
Db 27548 GCTGTGAACCAAGACGCGAGTATTAT 27524

RESULT 8
US-10-889-447-4/c
; Sequence 4, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; PRIOR FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 29291
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ02
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-4
Query Match 100.0%; Score 25; DB 19; Length 29291;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25
|||||
Db 27916 GCTGTGAACCAAGACGCGAGTATTAT 27892

RESULT 9
US-10-889-447-7/c
; Sequence 7, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; PRIOR FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 29430
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate GZ01
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-7
Query Match 100.0%; Score 25; DB 19; Length 29430;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25
|||||
Db 28051 GCTGTGAACCAAGACGCGAGTATTAT 28027

RESULT 10
US-10-839-729-15/c
; Sequence 15, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
```

; TITLE OF INVENTION: AND METHODS OF USE  
; FILE REFERENCE: BIOBANK.013A  
; CURRENT APPLICATION NUMBER: US/10/839,729  
; CURRENT FILING DATE: 2004-05-04  
; PRIOR APPLICATION NUMBER: 60/468703  
; PRIOR FILING DATE: 2003-05-06  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 29727  
; TYPE: DNA  
; ORGANISM: SARS Coronavirus  
US-10-839-729-15

Query Match 100.0%; Score 25; DB 18; Length 29727;  
Best Local Similarity 100.0%; Pred. No. 0.061;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25  
|||||  
DB 28286 GCTGTGAACCAAGACGCGAGTATTAT 28262

## RESULT 11

US-10-827-757-1/c  
; Sequence 1, Application US/10827757  
; Publication No. US20050004071A1  
; GENERAL INFORMATION:  
; APPLICANT: Comper, Wayne  
; TITLE OF INVENTION: Charged Polysaccharides Resistant To Lysosomal Degradation During  
; TITLE OF INVENTION: Kidney Filtration And Renal Passage And Their Use To Treat Or  
; TITLE OF INVENTION: Prevent Infection By Coronaviruses  
; FILE REFERENCE: 11213-007-999  
; CURRENT APPLICATION NUMBER: US/10/827,757  
; CURRENT FILING DATE: 2004-04-20  
; PRIOR APPLICATION NUMBER: 60/464,294  
; PRIOR FILING DATE: 2003-04-21  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 29727  
; TYPE: DNA  
; ORGANISM: SARS-related coronavirus (Urbani strain)  
US-10-827-757-1

Query Match 100.0%; Score 25; DB 18; Length 29727;  
Best Local Similarity 100.0%; Pred. No. 0.061;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25  
|||||  
DB 28286 GCTGTGAACCAAGACGCGAGTATTAT 28262

## RESULT 12

US-10-889-447-8/c  
; Sequence 8, Application US/10889447  
; Publication No. US20050075307A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Jain, Ravi  
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
; FILE REFERENCE: RTS-0685US  
; CURRENT APPLICATION NUMBER: US/10/889,447  
; CURRENT FILING DATE: 2004-07-12  
; PRIOR APPLICATION NUMBER: 60/486,670  
; PRIOR FILING DATE: 2003-07-12  
; NUMBER OF SEQ ID NOS: 241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 29727  
; TYPE: DNA  
; ORGANISM: SARS coronavirus Urbani

## US-10-889-447-8

Query Match 100.0%; Score 25; DB 19; Length 29727;  
Best Local Similarity 100.0%; Pred. No. 0.061;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25  
|||||  
DB 28286 GCTGTGAACCAAGACGCGAGTATTAT 28262

## RESULT 13

US-10-699-936-1/c  
; Sequence 1, Application US/10699936  
; Publication No. US20050095582A1  
; GENERAL INFORMATION:  
; APPLICANT: Gillim-Ross, Laura  
; APPLICANT: Taylor, Jill  
; APPLICANT: Scholl, David R.  
; APPLICANT: Wentworth, David E.  
; APPLICANT: Joillick, Joseph D.  
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory  
; TITLE OF INVENTION: Syndrome Coronavirus  
; FILE REFERENCE: DHI-07986  
; CURRENT APPLICATION NUMBER: US/10/699,936  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 29727  
; TYPE: DNA  
; ORGANISM: SARS coronavirus Urbani  
US-10-699-936-1

Query Match 100.0%; Score 25; DB 19; Length 29727;  
Best Local Similarity 100.0%; Pred. No. 0.061;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25  
|||||  
DB 28286 GCTGTGAACCAAGACGCGAGTATTAT 28262

## RESULT 14

US-10-839-729-17/c  
; Sequence 17, Application US/10839729  
; Publication No. US20050002953A1  
; GENERAL INFORMATION:  
; APPLICANT: Jens Herold  
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES  
; TITLE OF INVENTION: AND METHODS OF USE  
; FILE REFERENCE: BIOBANK.013A  
; CURRENT APPLICATION NUMBER: US/10/839,729  
; CURRENT FILING DATE: 2004-05-04  
; PRIOR APPLICATION NUMBER: 60/468703  
; PRIOR FILING DATE: 2003-05-06  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 29736  
; TYPE: DNA  
; ORGANISM: SARS Coronavirus  
US-10-839-729-17

Query Match 100.0%; Score 25; DB 18; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 0.061;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTAT 25  
|||||  
DB 28271 GCTGTGAACCAAGACGCGAGTATTAT 28247

RESULT 15

US-10-889-447-9/c  
 ; Sequence 9, Application US/10889447  
 ; Publication No. US20050075307A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bennett, C. Frank  
 ; APPLICANT: Jain, Ravi  
 ; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
 ; FILE REFERENCE: RTS-0685US  
 ; CURRENT APPLICATION NUMBER: US/10/889,447  
 ; CURRENT FILING DATE: 2004-07-12  
 ; PRIOR APPLICATION NUMBER: 60/486,670  
 ; PRIOR FILING DATE: 2003-07-12  
 ; NUMBER OF SEQ ID NOS: 241  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 29736  
 ; TYPE: DNA  
 ; ORGANISM: SARS coronavirus CUHK-W1  
 US-10-889-447-9

Query Match 100.0%; Score 25; DB 19; Length 29736;  
 Best Local Similarity 100.0%; Pred. No. 0.061;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGATATTAT 25  
 |||||  
 Db 28271 GCTGTGAACCAAGACGCGATATTAT 28247

Search completed: May 16, 2005, 14:33:50  
 Job time : 1186.16 secs

**THIS PAGE BLANK (uspto)**



```

RESULT 2
CK302404/c
LOCUS
DEFINITION
CK302404
SB02015B1H11.f1 normalized Keck-Tagu Library SB02 Taeniopygia
guttata cDNA clone SB02015B1H11.f1 5, mRNA sequence.
ACCESSION
CK302404
VERSION
CK302404.1 GI:44811978
SOURCE
EST.
Taeniopygia guttata
Taeniopygia guttata
Taeniopygia guttata
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archoosauria; Aves; Neognathae; Passeriformes; Estrildidae;
Estrildinae; Taeniopygia.
1 (bases 1 to 752)
REFERENCE
AUTHORS
Clayton,D.P., Arnold,A.P., Ball,G.F., Brenowitz,E., George,J.M.,
Mello,C.V., Wade,J., Replogle,K., Lewin,H., Band,M., Hernandez,A.
and Liu,L.
The Songbird Neurogenomics Initiative: An Evolving Public Resource
for Study of Genes, Brain, and Behavior
Unpublished (2004)
CONTACT: David F. Clayton
University of Illinois
B107 C/SL, 601 S. Goodwin, Urbana, IL 61801, USA
Tel: 217 244 3668
Fax: 217 244 1648
Email: dclayton@uiuc.edu
Base Calling/Quality Scores: PHRED from Washington University
Genome Center.
Vector Trimming: Cross match from Washington University Genome
Center PHRAP suite. Low quality bases (Phred score < 20) were
trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length. Funded
by PHS grant # RO1 NS045264, 'Songbird Neurogenomics Initiative.'
PCR Primers
FORWARD: TAATACGACTACTATAGGG(T7)
BACKWARD: ATTAACCTCTACTAAG(T3)
Insert Length: 752 Std Error: 0.00
Plate: SB02015B1 row: H column: 11
Seq primer: TAATACGACTACTATAGGG (T7)
High quality sequence stop: 752.
Location/Qualifiers
1. .752
/organism="Taeniopygia guttata"
/mol_type="mRNA"
/db_xref="taxon:59729"
/clone="SB02015B1H11.f1"
/tissue_type="brain"
/dev_stage="late embryo, post-hatch days 1, 10, 20, 45,
and adult (pooled)"
/lab_host="DH10B"
/clone_lib="normalized Keck-Tagu Library SB02"
/note="Organ: brain; Vector: pBS II SK(+); Site 1:
EcoRI(5' side of insert); Site 2: NotI (3' side of
insert); The library was constructed and normalized as
described by Bonaldo, M.F., Lennon, G. and Soares, M.B.
(1996), Genome Research 6(9): 791-806. An identifying tag
was added at the 3'during cDNA synthesis:
insertAAAAAAAAAAAAAAAAAATCGCA."

ORIGIN
Query Match 79.2%; Score 19.8; DB 7; Length 752;
Best Local Similarity 91.3%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TGTGACCAAGCGCAGTATTAT 25
|||||
Db 243 TGTGAACCAAGGCGAGATTAT 221
|||||

RESULT 3
BQ590872
LOCUS
DEFINITION
BQ590872
E012599-024-018-D14-SP6 MP1Z-ADIS-024-storage root Beta vulgaris
cDNA clone 024-018-D14 5-PRIME, mRNA sequence.
ACCESSION
BQ590872
VERSION
BQ590872.1 GI:26120455
SOURCE
EST.
Beta vulgaris
Beta vulgaris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Beta.
1 (bases 1 to 955)
REFERENCE
AUTHORS
Herwig,R., Schulz,B., Weisshaar,B., Hennig,S., Steinfath,M.,
Drungowski,M., Stahl,D., Wruck,W., Menze,A., O'Brien,J., Lehrach,H.
and Radelof,U.
Construction of a 'unigene' cDNA clone set by oligonucleotide
fingerprinting allows access to 25 000 potential sugar beet genes
Plant J. 32 (5), 845-857 (2002)
22362189
12472698
Contact: Weisshaar B
ADIS DNA core facility at MPIZ
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@piz-koeln.mpg.de
Insert Length: 955 Std Error: 0.00
Plate: 18 row: D column: 14
Seq primer: SP6; CATACGATTAGTGACACTATAG.
Location/Qualifiers
1. .955
/organism="Beta vulgaris"
/mol_type="mRNA"
/cultivar="KWS2320 (double haploid, monogerm breeding
line)"
/db_xref="taxon:161934"
/db_xref="GABI:189248"
/clone="024-018-D14"
/tissue_type="storage root"
/lab_host="EMDH10B"
/clone_lib="MP1Z-ADIS-024-storage root"
/note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;
cDNA library from sugar beet, library provided by KWS
Kleinwanzlebener Saatzzucht AG Einbeck, Germany, contact:
b.schulz@kws.de; cloning sites SalI-NotI, primer sites and
orientation:
SP6-Sali-CCACGCTCCG-5prime-cDNA-polyA-CC-NotI-T7; Note:
Sequencing granted in the context of the GABI-Beet
project, local PI: Dr. Katharina Schneider, coordinator:
Prof. Christian Jung; Sequence submission managed by
RZPD/GABI-Primary database: http://gabi.rzpd.de"

ORIGIN
Query Match 76.8%; Score 19.2; DB 5; Length 955;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGCGCAGTATTA 24
|||||
Db 53 GCTGTGAACCATGATGCAGTATGA 76
|||||

RESULT 4
AZ850688
LOCUS
DEFINITION
AZ850688
2M0152F11R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0152F11 R, genomic survey sequence.
ACCESSION
AZ850688
VERSION
AZ850688.1 GI:13035936
SOURCE
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



REFERENCE  
AUTHORS  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Irlam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D. Weiss, R.

TITLE  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL  
COMMENT  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0152 row: F column: 11  
Seq primer: CACACAGGAACACGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 330.

FEATURES  
source  
1..330  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0152F11"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (GI4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid r1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 74.4%; Score 18.6; DB 8; Length 330;  
Best Local Similarity 84.0%; Pred. No. 4.6e+02;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAACGACGATTTAT 25  
|||||  
Db 301 GCTGTGAACCAACGATTTAT 325

RESULT 5  
CG853284/c  
LOCUS  
DEFINITION  
ZMMBBb0351K10.r ZMMBBb Zea mays genomic clone ZMMBBb0351K10 3', genomic survey sequence.  
ACCESSION  
CG853284  
VERSION  
CG853284.1 GI:38380145  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
Zea mays  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 479)  
Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J. and Wing, R.  
Sequencing of the maize genome  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: T7  
BACKWARD: M13r  
Plate: 0351 row: K column: 10  
Seq primer: M13r  
Class: BAC ends.  
Location/Qualifiers  
1..479  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBBb0351K10"  
/lab\_host="DH10B"  
/clone\_lib="ZMMBBb"  
/notes="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2: HindIII; Zea mays L. spp. mays"

FEATURES  
source  
1..479  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBBb0351K10"  
/lab\_host="DH10B"  
/clone\_lib="ZMMBBb"  
/notes="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2: HindIII; Zea mays L. spp. mays"

ORIGIN  
Query Match 74.4%; Score 18.6; DB 9; Length 479;  
Best Local Similarity 84.0%; Pred. No. 4.8e+02;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAACGACGATTTAT 25  
|||||  
Db 475 GCTATGACCTAGACAGGATTTAT 451

RESULT 6  
CN781966  
LOCUS  
DEFINITION  
EST00061 cqlfloral Chenopodium quinoa cDNA clone F01C13 5' similar to 40S ribosomal protein S26 A2g40590, mRNA sequence.  
ACCESSION  
CN781966  
VERSION  
CN781966.1 GI:47561430  
KEYWORDS  
EST.  
SOURCE  
Chenopodium quinoa (quinoa)  
ORGANISM  
Chenopodium quinoa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Anaranthaceae; Chenopodium.

REFERENCE  
AUTHORS  
Coleman, C.E., Maughan, P.J., Stevens, M.R., Jellen, E.N. and Fairbanks, D.J.

TITLE  
Single nucleotide polymorphism discovery using quinoa expressed sequence tags  
JOURNAL  
COMMENT  
Unpublished (2004)  
Other ESTs: EST00098  
Contact: Coleman, Craig E.  
Department of Plant and Animal Sciences  
Brigham Young University  
275 WIDB, Brigham Young University, Provo, UT 84602, USA  
Tel: (801) 422-5145  
Fax: (801) 422-0008  
Email: craig\_coleman@byu.edu  
Plate: 01 row: C column: 13  
Seq primer: M13 Forward  
High quality sequence stop: 614.

```

FEATURES
source
Location/Qualifiers
1. .614
/organism="Chenopodium quinoa"
/mol_type="mRNA"
/cultivar="Real"
/db_xref="taxon:63459"
/clone="F01C13"
/dev_stage="Pre-anthesis"
/lab_host="XL-1 Blue"
/clone_lib="cgloral"
/note="Organ: Flowers; Vector: pTriplEx2; Site_1: SfiI; Site_2: SfiI; Floral cDNA Library from Chenopodium quinoa"

ORIGIN
Query Match 74.4%; Score 18.6; DB 7; Length 614;
Best Local Similarity 84.0%; Pred. No. 4.9e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGACGATATTAT 25
|||||
Db 229 GCTGTACACCAAGATCGAGTATTGT 253

RESULT 7
AZ122567
LOCUS RPCI-23-28P12 626 bp DNA linear GSS 12-MAY-2000
DEFINITION RPCI-23-28P12.TJ RPCI-23 Mus musculus genomic clone RPCI-23-28P12, genomic survey sequence.
ACCESSION AZ122567
VERSION AZ122567.1 GI:7790515
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 626)
Zhao,S., Nierman,W., Feildiyum,T., Malek,J., Shatsman,S., Akinret,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSSs: RPCI-23-28P12.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/cdb/bac ends/mouse/bac_end_intro.html
Plate: 28 row: P column: 12
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRV. Size selected DNA was cloned into the pBACe3.6 vector at the

FEATURES
source
Location/Qualifiers
1. .626
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-28P12"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse
```

LOCUS CN782003 672 bp mRNA linear EST 21-MAY-2004  
 DEFINITION EST00098 cgeed Chenopodium quinoa cDNA clone S01A15 5', similar to  
 40S ribosomal protein S26 At2g40590, mRNA sequence.  
 ACCESSION CN782003  
 VERSION CN782003.1 GI:47561467  
 KEYWORDS EST.  
 SOURCE Chenopodium quinoa (quinoa)  
 ORGANISM Chenopodium quinoa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllales; Amaranthaceae; Chenopodium.  
 REFERENCE 1 (bases 1 to 672)  
 AUTHORS Colanan,C.E., Maughan,P.J., Stevens,M.R., Jellen,E.N. and  
 Fairbanks,D.J.  
 TITLE Single nucleotide polymorphism discovery using quinoa expressed  
 sequence tags  
 JOURNAL Unpublished (2004)  
 COMMENT Other ESTs: EST00061  
 Contact: Coleman, Craig E.  
 Department of Plant and Animal Sciences  
 Brigham Young University  
 275 WIDB, Brigham Young University, Provo, UT 84602, USA  
 Tel: (801) 422-5145  
 Fax: (801) 422-0008  
 Email: craig\_coleman@byu.edu  
 Plate: 01 row: A column: 15  
 Seq primer: M13 Forward  
 High quality sequence stop: 672.  
 FEATURES  
 source Location/Qualifiers  
 1..672  
 /organism="Chenopodium quinoa"  
 /mol\_type="mRNA"  
 /cultivar="Real"  
 /db\_xref="taxon:63459"  
 /clone="S01A15"  
 /tissue\_type="Developing Seed"  
 /lab\_host="XL-1 Blue"  
 /clone\_lib="cgseed"  
 /notes="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI;  
 Developing Seed cDNA Library from Chenopodium quinoa"  
 ORIGIN  
 Query Match 74.4%; Score 18.6; DB 7; Length 672;  
 Best Local Similarity 84.0%; Pred. No. 4.9e+02;  
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 GCTGTGACCAAGACGCGATTAT 25  
 DB 259 GCTGTACACCAAGATGCGATTGT 283  
 RESULT 10  
 LOCUS CG137686/c 842 bp DNA linear GSS 21-AUG-2003  
 DEFINITION PUKBP30TB ZM 0.6-1.0 KB Zea mays genomic clone ZMMBTa0788G15,  
 genomic survey sequence.  
 ACCESSION CG137686  
 VERSION CG137686.1 GI:34028361  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 842)  
 AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
 Bennetzen,J.  
 TITLE Maize Genomics Consortium  
 JOURNAL Unpublished (2003)  
 COMMENT Other\_GSSs: PUKBP30TD  
 Contact: Cathy Whitelaw  
 Seq primer: TR  
 Class: sheared ends.  
 FEATURES  
 source Location/Qualifiers  
 1..945  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBTa0782E12"  
 /clone\_lib="ZM 0.6-1.0 KB"  
 /notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
 Cot selected genomic DNA library"  
 ORIGIN  
 Query Match 74.4%; Score 18.6; DB 9; Length 945;  
 Best Local Similarity 84.0%; Pred. No. 5.1e+02;  
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 GCTGTGACCAAGACGCGATTAT 25  
 DB 1 GCTGTGACCAAGACGCGATTAT 25

9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TF  
 Class: sheared ends.  
 FEATURES  
 source Location/Qualifiers  
 1..842  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBTa0788G15"  
 /clone\_lib="ZM 0.6-1.0 KB"  
 /notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
 Cot selected genomic DNA library"  
 ORIGIN  
 Query Match 74.4%; Score 18.6; DB 9; Length 842;  
 Best Local Similarity 84.0%; Pred. No. 5e+02;  
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 GCTGTGACCAAGACGCGATTAT 25  
 DB 56 GCTATGAACCTAGACAGGATTAT 32  
 RESULT 11  
 LOCUS CG116875 945 bp DNA linear GSS 20-AUG-2003  
 DEFINITION PUKBP30TB ZM 0.6-1.0 KB Zea mays genomic clone ZMMBTa0782E12,  
 genomic survey sequence.  
 ACCESSION CG116875  
 VERSION CG116875.1 GI:34000312  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 945)  
 AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
 Bennetzen,J.  
 TITLE Maize Genomics Consortium  
 JOURNAL Unpublished (2003)  
 COMMENT Other\_GSSs: PUKBP30TD  
 Contact: Cathy Whitelaw  
 Seq primer: TR  
 Class: sheared ends.  
 FEATURES  
 source Location/Qualifiers  
 1..945  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBTa0782E12"  
 /clone\_lib="ZM 0.6-1.0 KB"  
 /notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
 Cot selected genomic DNA library"  
 ORIGIN  
 Query Match 74.4%; Score 18.6; DB 9; Length 945;  
 Best Local Similarity 84.0%; Pred. No. 5.1e+02;  
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 GCTGTGACCAAGACGCGATTAT 25  
 DB 1 GCTGTGACCAAGACGCGATTAT 25

```

Db      454 GCTATGACCTAGACACAGCATTAT 478

RESULT 12
CN552254
LOCUS
DEFINITION
tael7c11.y1 Hydra EST Darmstadt I Hydra magnipapillata cDNA 5',
similar to SW:FMUH_HUMAN P07954 FUMARATE HYDRATASE, MITOCHONDRIAL
PRECUSOR ;, mRNA sequence.
ACCESSION
CN552254
VERSION
CN522254.1 GI:46961558
KEYWORDS
EST.
ORGANISM
Hydra magnipapillata
Eukaryota; Metazoa; Chnidaria; Hydrozoa; Hydroida; Anthomedusae;
Hydridae; Hydra.
REFERENCE
1 (bases 1 to 607)
AUTHORS
Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q.,
Martinez,D., Kibler,D., Hampson,S., Clifton,S., Pape,D., Marra,M.,
Hillier,L., Martin,J., Wyllie,T., Dancie,M., Theising,B., Bowers,Y.,
Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Teagareishvili,R.,
Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.
WashU Hydra EST Project
Unpublished (2002)
OTHER_ESTS: tael7c11.x1
Contact: Hans Bode
WashU Hydra EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library was constructed by Corina Guder / GATC Konstanz, Germany
Library materials provided by Thomas Holstein / Molecular Cell
Biology, TUD, Darmstadt DNA sequencing by: Washington University
Genome Sequencing Center For information on obtaining a clone
please contact: Hans Bode (hrobode@uci.edu)
Seq primer: -40UP
High quality sequence stop: 587.
FEATURES
source
1..607
/organism="Hydra magnipapillata"
/mol_type="mRNA"
/strain="sf-1 mutant of strain 105"
/db_xref="taxon:6085"
/lab_host="TransforMax EC100 (Epicentre), T1 Phage
resistant electrocompetent cells"
/clone_lib="Hydra EST Darmstadt I"
/note="Vector: pBluescript II SK (+); Site_1: NotI;
Site_2: EcoRI"
ORIGIN
Query Match 73.6%; Score 18.4; DB 7; Length 607;
Best Local Similarity 90.5%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 TGAACCAAGACGACGATTAT 25
||||| ||||| ||||| |||||
Db 433 TGAACCAAGACGACGATTAT 453

RESULT 13
AZ999859
LOCUS
DEFINITION
2M0287F19F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0287F19 F, genomic survey sequence.
ACCESSION
AZ999859
VERSION
AZ999859.1 GI:13871086
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Db      454 GCTATGACCTAGACACAGCATTAT 478

REFERENCE
1 (bases 1 to 728)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0287 row: F column: 19
Seq primer: CGTTGTAACGACGCGCCAGT
Class: Plasmid ends
High quality sequence stop: 728.
FEATURES
source
1..728
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0287F19"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN
Query Match 73.6%; Score 18.4; DB 8; Length 728;
Best Local Similarity 95.0%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGT 20
||||| ||||| ||||| |||||
Db 630 GCTGTGAACCAAGACAGT 649

RESULT 14
AW275602/c
LOCUS
DEFINITION
xp38b02.x1 NCI_CGAP_HN11 Homo sapiens cDNA clone IMAGE:2742603 3',
mRNA sequence.
ACCESSION
AW275602
VERSION
AW275602.1 GI:6662632
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS      1 (bases 1 to 187)
TITLE        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT      Tumor Gene Index
              Unpublished (1997)
              Contact: Robert Strausberg, Ph.D.
              Email: cgaps-r@mail.nih.gov
              Tissue Procurement: John F. Ensley, M.D., Silvio Gutkind Ph.D.,
              Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.
              cDNA Library Preparation: David B. Krizman, Ph.D.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html
              Seq primer: -40UP from Gibco
              High quality sequence stop: 186.
FEATURES     Location/Qualifiers
source       1..187
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:2742603"
              /tissue_type="normal squamous epithelium"
              /lab_host="DH10B"
              /clone_lib="NCI CGAP HN11"
              /notes="Organ: tongue; Vector: pAMP10; cDNA made by
              oligo-dT priming. Non-directionally cloned into the UDG
              sites of pAMP10. Size-selected on agarose gel, average
              insert size 500 bp. Primary library; non-amplified.
              cDNA Library Preparation: David B. Krizman, Ph.D (NCI).
              Reference: Krizman et al. (1996) Cancer Research
              56:5380-5383."
ORIGIN
Query Match      72.8%; Score 18.2; DB 2; Length 187;
Best Local Similarity 83.3%; Pred.No. 6.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTGTGAACCAAGACGCGAGTATTA 24
    |||||
DB 171 GATGTGAACCAANACAGTATTA 148

RESULT 15
245641/c
LOCUS      245641
DEFINITION HSCZQD031 normalized infant brain cDNA Homo sapiens cDNA clone
            c-zqd03, mRNA sequence.
ACCESSION 245641
VERSION   245641.1 GI:574873
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 289)
AUTHORS   Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C.,
            Devignes,M.D., Duprat,S., Houligatte,R., Jumeau,M.N., Lamy,B.,
            Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
            IMAGE: molecular integration of the analysis of the human genome
            and its expression
JOURNAL   C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
MEDLINE   95277534
PUBMED    7757816
COMMENT   Contact: Genethon
            Genexpress-Genethon
            Genethon Centre de recherche sur le Genome Humain
            1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
            Tel: 33169472800
            Fax: 33160778698
            Email: genexpress@genethon.fr

```

---

```

Single read.
Genexpress_library_id: C; Genexpress_sequence_id: y1c-zqd03
Seq primer: (-21)M13 universal.
FEATURES     Location/Qualifiers
source       1..289
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="c-zqd03"
              /sex="Female"
              /tissue_type="total brain"
              /dev_stage="3 months old"
              /clone_lib="normalized infant brain cDNA"
              /notes="Organ: brain; Vector: lafmid BA; Site 1: HindIII;
              Site 2: NotI; sex:Female; dev stage=3 months old;
              isolate=muscular atrophy patient; tissue type=total
              brain; total mRNA was oligo-(dT) primed and directionally
              cloned 5' -> 3' into the HindIII -> NotI sites of the
              lafmid BA vector. Clone library from B.Souares, Psychiatry
              Dept. Columbia University, USA. Normalization method:
              Bento Soares, P.N.A.S in press"
ORIGIN
Query Match      72.8%; Score 18.2; DB 7; Length 289;
Best Local Similarity 87.0%; Pred.No. 7.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CTGTGAACCAAGACGCGAGTATTA 24
    |||||
DB 246 CTGTGCACCAGGACTCAGTATTA 224

Search completed: May 16, 2005, 06:03:44
Job time : 2831.44 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 21:16:45 ; Search time 382.5 Seconds  
(without alignments)  
2026.885 Million cell updates/sec

Title: US-10-808-187A-2476  
Perfect score: 16  
Sequence: 1 accccaaggtttacc 16

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_da.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	1269	14 AY307165	AY307165 SARS coro
2	16	100.0	1269	14 AY360146	AY360146 SARS coro
3	16	100.0	1269	14 AY365036	AY365036 SARS coro
4	16	100.0	1669	14 AY536760	AY536760 SARS coro
5	16	100.0	1873	14 AY534758S4	AY534761 SARS coro
6	16	100.0	1938	14 AY534762S6	AY534767 SARS coro
7	16	100.0	2304	14 AY322205S4	AY322208 SARS coro
8	16	100.0	2810	14 AY290752	AY290752 SARS coro
9	16	100.0	3080	14 AY443086S10	AY443095 SARS coro
10	16	100.0	8439	14 AY304489	AY304489 SARS coro
11	16	100.0	8581	14 AY304487	AY304487 SARS coro
12	16	100.0	11006	14 AY304491	AY304491 SARS coro
13	16	100.0	11010	14 AY304493	AY304493 SARS coro
14	16	100.0	11010	14 AY304494	AY304494 SARS coro
15	16	100.0	13471	14 AY304490	AY304490 SARS coro
16	16	100.0	13471	14 AY304492	AY304492 SARS coro
17	16	100.0	29013	14 AY463060	AY463060 SARS coro
18	16	100.0	29350	14 AY394999	AY394999 SARS coro
19	16	100.0	29350	14 AY395000	AY395000 SARS coro

20	16	100.0	29350	14 AY395001	AY395001 SARS coro
21	16	100.0	29350	14 AY395002	AY395002 SARS coro
22	16	100.0	29433	14 AY394977	AY394977 SARS coro
23	16	100.0	29530	14 AY394985	AY394985 SARS coro
24	16	100.0	29573	14 AY338174	AY338174 SARS coro
25	16	100.0	29573	14 AY338175	AY338175 SARS coro
26	16	100.0	29573	14 AY348314	AY348314 SARS coro
27	16	100.0	29577	14 AY559094	AY559094 SARS coro
28	16	100.0	29592	14 AY463059	AY463059 SARS coro
29	16	100.0	29620	14 AY395004	AY395004 SARS coro
30	16	100.0	29640	14 AY394978	AY394978 SARS coro
31	16	100.0	29645	14 AY394979	AY394979 SARS coro
32	16	100.0	29646	14 AY394982	AY394982 SARS coro
33	16	100.0	29647	14 AY395003	AY395003 SARS coro
34	16	100.0	29661	14 AY559086	AY559086 SARS coro
35	16	100.0	29665	14 AY394988	AY394988 SARS coro
36	16	100.0	29670	14 AY559082	AY559082 SARS coro
37	16	100.0	29683	14 AY394996	AY394996 SARS coro
38	16	100.0	29683	14 AY394997	AY394997 SARS coro
39	16	100.0	29699	14 AY394983	AY394983 SARS coro
40	16	100.0	29705	14 AY283795	AY283795 SARS coro
41	16	100.0	29705	14 AY394980	AY394980 SARS coro
42	16	100.0	29706	14 AY283797	AY283797 SARS coro
43	16	100.0	29709	14 AY394987	AY394987 SARS coro
44	16	100.0	29710	14 AY559091	AY559091 SARS coro
45	16	100.0	29711	14 AY283794	AY283794 SARS coro

ALIGNMENTS

RESULT 1  
AY307165  
LOCUS 1269 bp RNA linear VRL 09-JUN-2003  
DEFINITION SARS coronavirus nucleocapsid protein (NP) gene, complete cds.  
ACCESSION AY307165  
VERSION AY307165.1 GI:31540948  
SOURCE .  
ORGANISM SARS coronavirus  
SARS coronavirus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
Coronaviridae; Coronavirus.  
1 (bases 1 to 1269)  
Sun, K., Anwar, A., Gupta, V., Tablin, M.T., Atkinson, R.,  
Chandrasekarn, A. and August, T.J.  
Direct Submission  
Submitted (26-MAY-2003) Genetic Immunotherapy Lab, Johns Hopkins  
Singapore, 41 Science Park Road, #03-18 The Gemini, Singapore,  
Singapore 117610, Singapore  
Location/Qualifiers  
1. .1269  
/organism="SARS coronavirus"  
/virion  
/mol\_type="genomic RNA"  
/db\_xref="taxon:227859"  
1. .1269  
/gene="NP"  
1. .1269  
/gene="NP"  
/notes="associated with virus genomic RNA; involved in CTL  
response"  
/product="nucleocapsid protein"  
/protein\_id="AAP49024.1"  
/db\_xref="GI:31540949"  
/translation="MSDNGPQSNQRAPRITFGGPTDSTDNNGNGRNGARPKORRPO  
KELSPNNTASFTALTQHKELRFPKQGVPTNPGDDQIGYRTRARRVGGDGM  
GIPNSPWFYVLTGTGPEALPYGANKGIVVWATEGALTPKDHITRNPNNAATVL  
OLPQGTTLPGKFGVAGSGSOASSRSSRSRSGNSRNSTPGSGRNSPARMSGGGET  
ALALLLDRLNOLSKVSKGQOQGTVTKKSAEASKKPROKRTATKQYNTQAFG  
RGPEPTQGNFGDQLIRQGTQYKHFQIAQFAPSASAFGMSRIGNEVTPSGTWLTY  
HGAIKLDDKDPQKFDKLVLLNKHDAIKTFPTPEPKKKKKTKDEAQPLPQKQKQPT  
VTLLPAADMDDFSRQLQNSMSGASDSTQA"

```

ORIGIN
Query Match      100.0%; Score 16; DB 14; Length 1269;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCAAGGTTTACCC 16
    |||||||
Db 126 ACCCAAGGTTTACCC 141

RESULT 2
AY360146
LOCUS      SARS coronavirus HPZ-2003 nucleocapsid protein (NP) gene, complete
DEFINITION
ACCESSION  AY360146
VERSION     AY360146.1 GI:34329618
KEYWORDS   SARS coronavirus HPZ-2003
SOURCE     SARS coronavirus HPZ-2003
ORGANISM   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE  1 (bases 1 to 1269)
AUTHORS    Zhu,H.P., Lu,Q.Y., Lu,Y.Y., Yao,P.P., Xu,F., Ge,Q., Weng,J.Q.,
            Yan,J.Y., Gong,L.M., Shi,W., Zhao,Z.Y. and Zhu,Z.Y.
TITLE      Cloning and expression of nucleocapsid protein gene of SARS
            associated coronavirus
JOURNAL    Unpublished
AUTHORS    2 (bases 1 to 1269)
            Zhu,H.P., Lu,Q.Y., Lu,Y.Y., Yao,P.P., Xu,F., Ge,Q., Weng,J.Q.,
            Yan,J.Y., Gong,L.M., Shi,W., Zhao,Z.Y. and Zhu,Z.Y.
TITLE      Direct Submission
JOURNAL    Submitted (04-AUG-2003) Zhejiang Provincial Key Laboratory of HFRS,
            Zhejiang Center for Disease Prevention and Control, 17 Lao Zhe Da
            Zhi Road, Hangzhou, Zhejiang 310009, P.R. China
FEATURES   source
            1..1269
                /organism="SARS coronavirus HPZ-2003"
                /mol_type="genomic RNA"
                /db_xref="taxon:243925"
            1..1269
                /gene="NP"
            1..1269
                /gene="NP"
                /codon_start=1
                /product="nucleocapsid protein"
                /protein_id="AAQ63890.1"
                /db_xref="GI:34329619"
                /translation="MSDNGPQSNRSAPRITFGGPTDSTDNNGRNGARPKQRPPQ
                GLPNNTASWFTALTQHGKEELRFRGGVPIINTNSGPDQDQIGYVRATRRVRGGDKM
                KELSPRWYFYLTGTPPEASLYPGANKGIVWVATEGALNTPKDHIGTRPNNNAATVL
                QLPQGTLLPKGFYAEGRSGSOASSRSRSGNSRSTPGSSRSGNSPARMASGGGET
                ALALLDLRLNLESKVSCKGQOQOQVTVTKSAEASKKPRKRTATKQYNVTOAFG
                RRGPEOTQNGFDQDLIRQGTDKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY
                HGAIKDDKDPQKDNVILNKHIDAYKTFPTEPKDKKKKTDQAQPLPQRKKQPT
                VTLLPAADMDDFSRLQNSMSGASADSTQA"

gene
CDS

ORIGIN
Query Match      100.0%; Score 16; DB 14; Length 1269;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCAAGGTTTACCC 16
    |||||||
Db 126 ACCCAAGGTTTACCC 141

RESULT 4
AY36760
LOCUS      SARS coronavirus BJ01 nucleocapsid protein mRNA, complete cds.
DEFINITION
ACCESSION  AY36760
VERSION     AY36760.3 GI:52100973
KEYWORDS   SARS coronavirus BJ01
SOURCE     SARS coronavirus BJ01
ORGANISM   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE  1 (bases 1 to 1669)
AUTHORS    Li,T., Li,X., Chang,Z. and Liu,L.
TITLE      Identification of SARS-CoV mRNA leader sequence
JOURNAL    Unpublished
AUTHORS    2 (bases 1 to 1669)
            Li,T., Li,X., Liu,L. and Chang,Z.
TITLE      Direct Submission
JOURNAL    Submitted (30-JAN-2004) Institute of Biomedicine,Tsinghua

ORIGIN
Query Match      100.0%; Score 16; DB 14; Length 1269;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCAAGGTTTACCC 16
    |||||||
Db 126 ACCCAAGGTTTACCC 141

RESULT 3
AY365036
LOCUS      SARS coronavirus HB nucleocapsid protein (NP) gene, complete cds.
DEFINITION
ACCESSION  AY365036
VERSION     AY365036.1 GI:38176101
KEYWORDS   SARS coronavirus HB
SOURCE     SARS coronavirus HB
ORGANISM   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE  1 (bases 1 to 1269)
AUTHORS    Timani,K.A., Ye,L., Ye,L., Ye,L., Ye,L., Zhu,Y. and Guo,D.
TITLE      Direct Submission
JOURNAL    Submitted (11-AUG-2003) Institute of Virology, Wuhan University,
            Wuhan, Hubei 430072, P.R. China
FEATURES   source
            1..1269
                /organism="SARS coronavirus HB"
                /mol_type="genomic RNA"
                /isolate="HB"
                /db_xref="taxon:253634"
                /country="China: Wuhan, Hubei"
            1..1269
                /gene="NP"
            1..1269
                /gene="NP"
                /note="N protein"
                /codon_start=1
                /product="nucleocapsid protein"
                /protein_id="AAR12990.1"
                /db_xref="GI:38176102"
                /translation="MSDNGPQSNRSAPRITFGGPTDSTDNNGRNGARPKQRPPQ
                GLPNNTASWFTALTQHGKEELRFRGGVPIINTNSGPDQDQIGYVRATRRVRGGDKM
                KELSPRWYFYLTGTPPEASLYPGANKGIVWVATEGALNTPKDHIGTRPNNNAATVL
                QLPQGTLLPKGFYAEGRSGSOASSRSRSGNSRSTPGSSRSGNSPARMASGGGET
                ALALLDLRLNLESKVSCKGQOQOQVTVTKSAEASKKPRKRTATKQYNVTOAFG
                RRGPEOTQNGFDQDLIRQGTDKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY
                HGAIKDDKDPQKDNVILNKHIDAYKTFPTEPKDKKKKTDQAQPLPQRKKQPT
                VTLLPAADMDDFSRLQNSMSGASADSTQA"

gene
CDS

ORIGIN
Query Match      100.0%; Score 16; DB 14; Length 1269;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCAAGGTTTACCC 16
    |||||||
Db 126 ACCCAAGGTTTACCC 141

RESULT 4
AY36760
LOCUS      SARS coronavirus BJ01 nucleocapsid protein mRNA, complete cds.
DEFINITION
ACCESSION  AY36760
VERSION     AY36760.3 GI:52100973
KEYWORDS   SARS coronavirus BJ01
SOURCE     SARS coronavirus BJ01
ORGANISM   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE  1 (bases 1 to 1669)
AUTHORS    Li,T., Li,X., Chang,Z. and Liu,L.
TITLE      Identification of SARS-CoV mRNA leader sequence
JOURNAL    Unpublished
AUTHORS    2 (bases 1 to 1669)
            Li,T., Li,X., Liu,L. and Chang,Z.
TITLE      Direct Submission
JOURNAL    Submitted (30-JAN-2004) Institute of Biomedicine,Tsinghua

```



```

REFERENCE
AUTHORS University, Beijing, Beijing 100084, China
3 (bases 1 to 1669)
TITLE Direct Submission
JOURNAL Submitted (06-JUL-2004) Institute of Biomedicine, Tsinghua
University, Beijing, Beijing 100084, China
REMARK Sequence update by submitter
AUTHORS 4 (bases 1 to 1669)
TITLE Direct Submission
JOURNAL Submitted (15-SEP-2004) Institute of Biomedicine, Tsinghua
University, Beijing 100084, China
REMARK Sequence update by submitter
COMMENT On Sep 15, 2004 this sequence version replaced gi:49921010.
FEATURES
    source
        location/Qualifiers
            1..1669
                /organism="SARS coronavirus BJ01"
                /mol_type="mRNA"
                /strain="BJ01"
                /db_xref="taxon:228407"
                /country="China"
                81..1349
                    /codon_start=1
                    /product="nucleocapsid protein"
                    /protein_id="AAS48456.1"
                    /db_xref="GI:44889746"
                    /translation="MSDNGPQSNORSAPRTFGGTDTSDNNQNGRNGARP KORRPQ
                    GLPNNTASFWTLTQHGKEELRFPQGVQVINTSGPDDQIGYRATRRVRGDDGKM
                    KSLSPWRYFYLTGPASLPYGANKGIVVWATGALNTPKDHIGTRPNNAATVTL
                    QLPQGTTLPGFYAEGSRGGSOASSRSSRSGNSRSTPGSSRSGNSPAPASGGGET
                    ALALLLLDLNQLSKSVSGKQQCGQTVTKSAEASKPKPKRKTATKYNTQTAFG
                    RRPEQTQGNFGQDLIRQGTIDYKHWPTQIAQFAPSAFFGMSRIGHEVTPSGTWLT
                    HGAIKDDKDPQKDNVLLNKHIDAYKTPTPEPKKKKKTKDEAQLPQPKKQPT
                    VTPLPADMDDFSRQLQNSGASADSTQA"
    CDS
        1..1669
            /organism="SARS coronavirus BJ01"
            /mol_type="mRNA"
            /strain="BJ01"
            /db_xref="taxon:228407"
            /country="China"
            81..1349
                /codon_start=1
                /product="nucleocapsid protein"
                /protein_id="AAS48456.1"
                /db_xref="GI:44889746"
                /translation="MSDNGPQSNORSAPRTFGGTDTSDNNQNGRNGARP KORRPQ
                GLPNNTASFWTLTQHGKEELRFPQGVQVINTSGPDDQIGYRATRRVRGDDGKM
                KSLSPWRYFYLTGPASLPYGANKGIVVWATGALNTPKDHIGTRPNNAATVTL
                QLPQGTTLPGFYAEGSRGGSOASSRSSRSGNSRSTPGSSRSGNSPAPASGGGET
                ALALLLLDLNQLSKSVSGKQQCGQTVTKSAEASKPKPKRKTATKYNTQTAFG
                RRPEQTQGNFGQDLIRQGTIDYKHWPTQIAQFAPSAFFGMSRIGHEVTPSGTWLT
                HGAIKDDKDPQKDNVLLNKHIDAYKTPTPEPKKKKKTKDEAQLPQPKKQPT
                VTPLPADMDDFSRQLQNSGASADSTQA"
    ORIGIN
        Query Match 100.0%; Score 16; DB 14; Length 1669;
        Best Local Similarity 100.0%; Pred. No. 2.9e+02;
        Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

        QY 1 ACCCCAGGTTTACCC 16
           |||||
        DB 206 ACCCCAGGTTTACCC 221
           |||||

        RESULT 5
        AY534758S4 1873 bp RNA linear VRL 17-MAR-2004
        LOCUS SARS coronavirus Sin0409, partial sequence.
        DEFINITION AY534761
        VERSION AY534761.1 GI:45384968
        KEYWORDS
        SEGMENT
        SOURCE
            ORGANISM
                SARS coronavirus Sin0409
                Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
                Coronaviridae; Coronavirus.
        REFERENCE
            1 (bases 1 to 1873)
            Lim, P.L., Kurup, A., Gopalakrishna, G., Chan, K.P., Wong, C.W.,
            Ng, L.C., Se-Thoe, S.Y., Oon, L., Bai, X., Stanton, L.W., Ruan, Y.,
            Miller, L.D., Vega, V.B., James, L., Ooi, P.L., Kai, C.S., Olsen, S.J.,
            Ang, B. and Leo, Y.S.
            Laboratory-acquired severe acute respiratory syndrome (SARS) -
            Singapore 2003
            Unpublished
        JOURNAL
        AUTHORS
        TITLE Direct Submission
        JOURNAL Submitted (28-JAN-2004) Genome Institute of Singapore, 60, Biopolis
        Street, 02-01, Genome, Singapore 138672, Singapore
        FEATURES
            source
                location/Qualifiers
                    1..1873
                        /organism="SARS coronavirus Sin0409"
                        /mol_type="genomic RNA"
                        /db_xref="taxon:266148"
                        /country="Singapore"

        ORIGIN
        Query Match 100.0%; Score 16; DB 14; Length 1873;
        Best Local Similarity 100.0%; Pred. No. 2.9e+02;
        Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

        QY 1 ACCCCAGGTTTACCC 16
           |||||
        DB 206 ACCCCAGGTTTACCC 221
           |||||

        RESULT 5
        AY534758S4 1873 bp RNA linear VRL 17-MAR-2004
        LOCUS SARS coronavirus Sin0409, partial sequence.
        DEFINITION AY534761
        VERSION AY534761.1 GI:45384968
        KEYWORDS
        SEGMENT
        SOURCE
            ORGANISM
                SARS coronavirus Sin0409
                Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
                Coronaviridae; Coronavirus.
        REFERENCE
            1 (bases 1 to 1873)
            Lim, P.L., Kurup, A., Gopalakrishna, G., Chan, K.P., Wong, C.W.,
            Ng, L.C., Se-Thoe, S.Y., Oon, L., Bai, X., Stanton, L.W., Ruan, Y.,
            Miller, L.D., Vega, V.B., James, L., Ooi, P.L., Kai, C.S., Olsen, S.J.,
            Ang, B. and Leo, Y.S.
            Laboratory-acquired severe acute respiratory syndrome (SARS) -
            Singapore 2003
            Unpublished
        JOURNAL
        AUTHORS
        TITLE Direct Submission
        JOURNAL Submitted (28-JAN-2004) Genome Institute of Singapore, 60, Biopolis
        Street, 02-01, Genome, Singapore 138672, Singapore
        FEATURES
            source
                location/Qualifiers
                    1..1873
                        /organism="SARS coronavirus Sin0409"
                        /mol_type="genomic RNA"
                        /db_xref="taxon:266148"
                        /country="Singapore"

        ORIGIN
        Query Match 100.0%; Score 16; DB 14; Length 1873;
        Best Local Similarity 100.0%; Pred. No. 2.9e+02;
        Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

        QY 1 ACCCCAGGTTTACCC 16
           |||||
        DB 471 ACCCCAGGTTTACCC 486
           |||||

        RESULT 7
        AY322205S4 2304 bp RNA linear VRL 21-JUL-2003
        LOCUS SARS coronavirus Shanghai LY Orf7a gene, partial cds; and Orf7b,
        DEFINITION AY322208
        VERSION AY322208.1 GI:32454342
        KEYWORDS
        SEGMENT
        SOURCE
            ORGANISM
                SARS coronavirus Shanghai LY
                Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
                Coronaviridae; Coronavirus.
        REFERENCE
            1 (bases 1 to 2304)
            Yuan, Z., Zhang, X., Hu, Y., Lan, S., Wang, H., Zhou, Z. and Wen, Y.

```

```

ORIGIN
    Query Match 100.0%; Score 16; DB 14; Length 1873;
    Best Local Similarity 100.0%; Pred. No. 2.9e+02;
    Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1 ACCCCAGGTTTACCC 16
       |||||
    DB 445 ACCCCAGGTTTACCC 460
       |||||

    RESULT 6
    AY534762S6 1938 bp RNA linear VRL 17-MAR-2004
    LOCUS SARS coronavirus Sin_WNV, partial sequence.
    DEFINITION AY534767
    ACCESSION AY534767.1 GI:45384975
    VERSION
    KEYWORDS
    SEGMENT
    SOURCE
        ORGANISM
            SARS coronavirus Sin_WNV
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
    REFERENCE
        1 (bases 1 to 1938)
        Lim, P.L., Kurup, A., Gopalakrishna, G., Chan, K.P., Wong, C.W.,
        Ng, L.C., Se-Thoe, S.Y., Oon, L., Bai, X., Stanton, L.W., Ruan, Y.,
        Miller, L.D., Vega, V.B., James, L., Ooi, P.L., Kai, C.S., Olsen, S.J.,
        Ang, B. and Leo, Y.S.
        Laboratory-acquired severe acute respiratory syndrome (SARS) -
        Singapore 2003
        Unpublished
    JOURNAL
    AUTHORS
    TITLE Direct Submission
    JOURNAL Submitted (28-JAN-2004) Genome Institute of Singapore, 60, Biopolis
    Street, 02-01, Genome, Singapore 138672, Singapore
    FEATURES
        source
            location/Qualifiers
                1..1938
                    /organism="SARS coronavirus Sin_WNV"
                    /mol_type="genomic RNA"
                    /db_xref="taxon:266148"
                    /country="Singapore"

    ORIGIN
    Query Match 100.0%; Score 16; DB 14; Length 1938;
    Best Local Similarity 100.0%; Pred. No. 2.9e+02;
    Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1 ACCCCAGGTTTACCC 16
       |||||
    DB 471 ACCCCAGGTTTACCC 486
       |||||

    RESULT 7
    AY322205S4 2304 bp RNA linear VRL 21-JUL-2003
    LOCUS SARS coronavirus Shanghai LY Orf7a gene, partial cds; and Orf7b,
    DEFINITION AY322208
    VERSION AY322208.1 GI:32454342
    KEYWORDS
    SEGMENT
    SOURCE
        ORGANISM
            SARS coronavirus Shanghai LY
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
    REFERENCE
        1 (bases 1 to 2304)
        Yuan, Z., Zhang, X., Hu, Y., Lan, S., Wang, H., Zhou, Z. and Wen, Y.

```

TITLE JOURNAL	Direct Submission Submitted (12-JUN-2003) Molecular Virology, Shanghai Medical College of Fudan University, 138 Yi Xue Yuan Road, Shanghai 200032, P.R. China
FEATURES source	Location/Qualifiers 1..2304 /organism="SARS coronavirus Shanghai LY" /mol_type="genomic RNA" /db_xref="taxon:235173" /country="China: Shanghai" <1..276 /codon_start=1 /product="Orf7a" /protein_id="AAP82970.1" /db_xref="GI:32454348" /translation="KEPCPSATYEGNSPFHPLADNKFKALCTCTSTHFAFACADGTRHTY OLRARSVGPLRQBEVQEQLYSLFLVLAALVFLILCFTIKRKE" 273..407 /codon_start=1 /product="Orf7b" /protein_id="AAP82971.1" /db_xref="GI:32454349" /translation="WNELTLIDFYLCFLAFLFLVLMILIFWFSLEIQDLEPCTKV "
CDS	414..533 /codon_start=1 /product="Orf8a" /protein_id="AAP82972.1" /db_xref="GI:32454350" /translation="MKLLIVLTCISLCSCICTVQRCASNKPHVLEDPCKVQH" 499..753 /codon_start=1 /product="Orf8b" /protein_id="AAP82973.1" /db_xref="GI:32454351" /translation="MCLKILVYTRGNTYSTAWLCALGKVLPEHRHWTVQTCFNV TINCODPAGGALLACWYLHEGHQIAAFRDVLVLNKRIN" 755..2023 /codon_start=1 /product="nucleocapsid protein" /protein_id="AAP82974.1" /db_xref="GI:32454352" /translation="MSDNGPQSNQSRAPRITFGGPTDSTDNNGRNGARPKORRPQ GLPNFTASFTALTGHKEELRPFGQVPINTNSGPDQIGIYRATRRVRRGGDKM KELSPRWYFYLGTPGEASLPYCANKEGIVWATEGALNPKDHIGTRPNNAATVL QLPQGTLLPKGYAEGSRGSGQSSRSRGRNRPSTPGSRGNSPARVASGGGET ALALLDLNOLSKVSGKQQOQOTVTKSAEASKPKRKTATKQYNVTOAFG RSGPQTCNFGDQDLIRGTDYKHPQIAQAPASAFEGMSRIGMEAPSGTWLTY HGAIKDDKDPQFKNVLLNKHIDAYKTFPPTEPKDKKKKTDEAQPLPQRQKQPT VTLLPAADMDDFSRQLQNSGASADSTQA"
ORIGIN	Query Match 100.0%; Score 16; DB 14; Length 2304; Best Local Similarity 100.0%; Pred No. 2.8e+02; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  QY 1 ACCCCAAGTTTACCC 16       Db 880 ACCCCAAGTTTACCC 895  RESULT 8 LOCUS AY290752 DEFINITION SARS coronavirus ZJ01 isolate ZJ01b uncharacterized protein 6, uncharacterized protein 7a, uncharacterized protein 7b, uncharacterized protein 8a, uncharacterized protein 8b, nucleocapsid protein, uncharacterized protein 9b, and uncharacterized protein 9c genes, complete cds. ACCESSION AY290752 VERSION AY290752.2 GI:38735509 KEYWORDS SARS coronavirus ZJ01
ORGANISM	SARS coronavirus ZJ01
REFERENCE AUTHORS	Viruses: sRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus. 1 (bases 1 to 2810) Li,L., Wang,Z., Lu,Y., Bao,Q., Chen,S., Wu,N., Cheng,S., Weng,J., Zhang,Y., Yan,J., Mei,L., Wang,X., Zhu,H., Yu,Y., Zhang,M., Li,M., Yao,J., Lu,Q., Yao,P., Bo,X., Wo,J., Wang,S. and Hu,S. Severe acute respiratory syndrome-associated coronavirus genotype and its characterization Chin. Med. J. 116 (9), 1288-1292 (2003) 22889812 PUBMED 14527350
TITLE	2 (bases 1 to 2810)
JOURNAL MEDLINE PUBMED	Wang,Z.G., Li,L.J., Luo,Y., Zhang,J.Y., Wang,M.Y., Cheng,S.Y., Zhang,J.J., Wang,X.M., Lu,Y., Wu,N.P., Mei,L.L. and Wang,Z.X. Molecular biological analysis of genotyping and phylogeny of severe acute respiratory syndrome associated coronavirus Chin. Med. J. 117 (1), 42-48 (2004) 14733771 PUBMED
REFERENCE AUTHORS	3 (bases 1 to 2810)
TITLE	Wang,Z., Cheng,S., Zhang,Y., Lu,Y. and Mei,L.
JOURNAL	Direct Submission Submitted (03-MAY-2003) Department of Microbiology, Zhejiang Center of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou, Zhejiang 310009, China 4 (bases 1 to 2810)
REFERENCE AUTHORS	Wang,Z., Cheng,S., Zhang,Y., Lu,Y. and Mei,L.
TITLE	Direct Submission
JOURNAL	Submitted (05-DEC-2003) Department of Microbiology, Zhejiang Center of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou, Zhejiang 310009, China Nucleotide and amino acid sequences updated by submitter On Dec 5, 2003 this sequence version replaced gi:31505969. Location/Qualifiers 1..2810 /organism="SARS coronavirus ZJ01" /mol_type="genomic RNA" /isolate="ZJ01b" /isolation_source="patient with severe acute respiratory syndrome" /db_xref="taxon:230471" /country="China: Hangzhou" 1..6 /note="transcription regulatory core sequence 6" 162..353 /codon_start=1 /product="uncharacterized protein 6" /protein_id="AAR27473.1" /db_xref="GI:38735510" /translation="MFHLVDVQVTAETLIIIMRTFRIWNLVDVIISIVRQLFKPL TKKNYSELDDDEPMELDYP" 331 /note="Xiaohong genotype locus 6" 355..360 /note="transcription regulatory core sequence 7" 361..729 /codon_start=1 /product="uncharacterized protein 7a" /protein_id="AAR27474.1" /db_xref="GI:38735511" /translation="WKIILFLTVFTSCELYHOCVGRGTTVLKKEPCPSCTYEGNS PFHPLADNKFKALCTCTSTHFAFACADGTRHTYQLRARSVSPKLFIRQBEVQEQLYSLFL LIVLAALVFLILCFTIKRKE" 726..860 /codon_start=1 /product="uncharacterized protein 7b" /protein_id="AAR27475.1" /db_xref="GI:38735512" /translation="WNEULTLIDFYLCFLAFLFLVLMILIFWFSLEIQDLEPCTKV "
misc_feature	861..866 /note="transcription regulatory core sequence 8"
CDS	867..986

```

/codon_start=1
/product="uncharacterized protein 8a"
/protein_id="AAR27476.1"
/db_xref="GI:38735513"
/translation="MKLLIVLTGICSLCICCTVQRCASKNKPHVLEDPCKVQH"
915
/misc_feature
952..1206
/codon_start=1
/product="uncharacterized protein 8b"
/protein_id="AAR27477.1"
/db_xref="GI:38735514"
/translation="MCLKILRVYNTGNTYSTAWLCAHGKVLPHFRHWHVQCTPNV
TINCQDPAGALLARCYLHEGHQTAAPRDVVLVNRKN"
1193..1198
/misc_feature
1208..2476
/notes="transcription regulatory core sequence 9"
/codon_start=1
/product="nucleocapsid protein"
/protein_id="AAP44772.1"
/db_xref="GI:31505970"
/translation="MSDNGPQSNQRSAPRITFGGPTDSTDNQNGRNGARPKQRPPQ
GLPNTASWFTALTQHGKELRPRGQGVPIINTSGDDOIGYVRRATRVRGDGM
KELSPRYFYLTGPEASLPYGANKEGIVVATEGALNPKDHICTRNPNNAATVL
QLPQGTLPKFPYAGSRGSGQASSRSSRSRSGNSRSTFGSRGNSPARMASGGSET
ALALLDLRLNQLSKVSGKQGGQGTQVTKSAEASKPRKRRTATKQYNTQAPG
RRPEQTQNGFDLIRQGTQYKHPQIAQFAPSASAFEGMSRIGMEVTPSGTWLTY
HGAIKLDDKDPQPKDNVILNKHIDAYKTPPTPEPKDKKKKTDEAQLPQRQKKQPT
VTLLPAADMDPSRQLQNSMGASADSTQA"
1218..1514
/codon_start=1
/product="uncharacterized protein 9b"
/protein_id="AAP69659.1"
/db_xref="GI:31747858"
/translation="MDPNQNVVPPALHLVDPOILTITRMDAMGQONSADPKVYP
IILRLGSQLSLSWARNLDSLEARFQSTIVVQMTKLATTEBELPDEFVVVTA"
1671..1883
/codon_start=1
/product="uncharacterized protein 9c"
/protein_id="AAP69660.1"
/db_xref="GI:31747859"
/translation="MLPPCYNFKLEQHCQKASTQREAEAAVKPLLAPHVAVIQEIQ
LLAAVEILLLEWIAEVVKLPSRYCC"

ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 2810;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCCAAGTTTATCCC 16
|||||
Db 1333 ACCCCAAGTTTATCCC 1348

RESULT 9
AY443086S10 3080 bp RNA linear VRL 02-FEB-2004
LOCUS SARS coronavirus CUHK-L2 orf4 and orf3 genes, partial cds; small
DEFINITION envelope E protein, membrane glycoprotein M, orf7, orf8, orf9,
orf10, and orf11 genes, complete cds; nucleocapsid protein gene,
partial cds; and orf13 gene, complete cds.
ACCESSION AY443095
VERSION AY443095.1 GI:41352885
KEYWORDS
SEGMENT 10 of 10
SOURCE SARS coronavirus CUHK-L2
ORGANISM SARS coronavirus CUHK-L2
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 3080)
AUTHORS Chim,S.S.C., Tong,Y.K., Hung,E.C.W., Chiu,R.W.K. and Lo,Y.M.D.
TITLE Genomic Sequencing of a SARS Coronavirus Isolate That Predated the
Metropole Hotel Case Cluster in Hong Kong
```

```

JOURNAL Clin. Chem. 50 (1), 231-233 (2004)
PUBMED 14709660
REFERENCE 2 (bases 1 to 3080)
AUTHORS Chim,S.S.C., Tong,Y.K., Hung,E.C., Chiu,R.W. and Lo,Y.M.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-2003) Chemical Pathology, Chinese University of
Hong Kong, Prince of Wales Hospital, 30-32 Ngan Shing Street,
Shatin, Hong Kong, China
FEATURES
Location/Qualifiers
source
1..3080
/organism="SARS coronavirus CUHK-L2"
/virus
/mol_type="genomic RNA"
/db_xref="taxon:260550"
<1..328
/codon_start=2
/product="orf4"
/protein_id="AAS01066.1"
/db_xref="GI:41352897"
/translation="VVILRIGTVLTKMSLYMAISPKFTTSLSLHKLQTLVLKQLHS
SSLTSLKTRMCKYQTQSTALQELLIQWIFQWMSRRRLACLCKKKVSTNLCTHSF
RKQQR"
<1..267
/codon_start=1
/product="orf3"
/protein_id="AAS01065.1"
/db_xref="GI:41352896"
/translation="GGYSEDRHSGVKDYVVVHGVTVEVYQLESTQITTDGTIENATF
FIPNKLVPKPPNVQIHTIDGSGVANPAMPDIYDEPTTTSVPL"
292..522
/codon_start=1
/product="small envelope E protein"
/protein_id="AAS01067.1"
/db_xref="GI:41352898"
/translation="MYSFVSEETGLIVNSVLFLAFVVELVTLALTALRLCAYCC
NIWVSLVAPTVVYVSRVKNLSSEGVDDL"
573..1238
/codon_start=1
/product="membrane glycoprotein M"
/protein_id="AAS01068.1"
/db_xref="GI:41352899"
/translation="MADNGTITVEELKOLLEONLVIIGELFLAWIMLQFAYSNRHF
LYIKLVFLLMPTVLAFLVLAAYTRINWVGIAIACIVGLMWLSYFVASFLRF
ATRTSMSPNPTNILLNPLRGITVTRPLMSELVIGAVIIRHRLRMAGHSIGRCDI
KDLPEITVATSRITLSYYKLIGASQRVGTDGFAAYNRYRIGNYKLTNDHAGSNDNIAL
LVO"
1249..1440
/codon_start=1
/product="orf7"
/protein_id="AAS01069.1"
/db_xref="GI:41352900"
/translation="MFHLVDQVTVAEILIIIMRTFRIAWNLDVIISIVRQLFKPL
TKKNYSELDDDEPMELDYP"
1448..1816
/codon_start=1
/product="orf8"
/protein_id="AAS01070.1"
/db_xref="GI:41352901"
/translation="MKIILFLTLIVTSCELYHQCVRGTTVLLKEPCPSGTVEGNS
PHPLADNRFALCTSTHFACADGTRHTYQURARSVSKLFIHQEEVQOGLYSPLF
LVAALVFLILCTIKRKE"
1813..1947
/codon_start=1
/product="orf9"
/protein_id="AAS01071.1"
/db_xref="GI:41352902"
/translation="MNELTLIDFVLCFLAFLFLVLIIMLIIFWFSLEIQDLPEPCTKV"
1954..2073
/codon_start=1
/product="orf10"
/protein_id="AAS01072.1"
/db_xref="GI:41352903"
```

```
/translation="MKLLIVLTCISLCIRTVVQRCASNKPHVLEDFCKVQH"
2039..2293
/codon_start=1
/product="orf11"
/protein_id="AAS01073.1"
/db_xref="GI:41352904"
/translation="WCLKILVRYNTRGNTYSTAWLCALGKVLPRHRWHTVQTCFNV
TINCQDPAGGALIARCWYLHEGHQTAAPRDVLVLNKRIN"
2295..>3080
CDS
/codon_start=1
/product="nucleocapsid protein"
/protein_id="AAS01074.1"
/db_xref="GI:41352905"
/translation="MSDNGPQSNQSRAPRIIFGGTSDTDNNQNGRNGARPQKRRPQ
GLPNNVAFWFTALTQHGKEELRFPQGQVPINTNSGDDQIGIYRATRRVRGGDKM
KELSPRWYFPLGTGPEASLPYGANKGIVWVATEXXXXXXXKXXXXXXXXXXXX
XLPQGTTPFKGYAEGSRGGSQASRSSRSRXXNSRNSTPGSSRSGNSPARMASGGGT
ALALLDLRLNQLESKVGSGQQQGXTVTKSAAEASKPRQK"
2305..2601
CDS
/codon_start=1
/product="orf13"
/protein_id="AAS01075.1"
/db_xref="GI:41352906"
/translation="MDPNOTVVPALHLVDPOIQIOTITRMEDAMGQONGADPKVYP
IILRGSQLSLSMARNDLSLEARAFQSTPIVVQMTKLATTEELPDEFVVVTAK"
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 3080;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACCC 16
| | | | | | | | | | | | | | | |
DB 2420 ACCCAAGGTTTACCC 2435

RESULT 10
AY304489
LOCUS SARS coronavirus SZ1, partial genome.
DEFINITION
ACCESSION AY304489
VERSION AY304489.1 GI:34482140
KEYWORDS
SOURCE SARS coronavirus SZ1
ORGANISM SARS coronavirus SZ1
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
AUTHORS Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L.,
Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., Butt, K.M., Wong, K.L.,
Chan, K.W., Lim, W., Shortridge, K.F., Yuen, K.Y., Peiris, J.S.M. and
Poon, L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 22913660
PubMed 12958366
REFERENCE
AUTHORS Guan, Y. and Zheng, B.J.
Direct Submission
TITLE Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
source
1..8439
/organism="SARS coronavirus SZ1"
/mol_type="genomic RNA"
/isolate="SZ1"
/db_xref="taxon:231516"
/country="Hong Kong"
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 8439;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACCC 16
| | | | | | | | | | | | | | | |
DB 7109 ACCCAAGGTTTACCC 7124

RESULT 12
AY304491
LOCUS SARS coronavirus GZ60, partial genome.
DEFINITION
ACCESSION AY304491
VERSION AY304491.1 GI:34482142
KEYWORDS
SOURCE SARS coronavirus GZ60
ORGANISM SARS coronavirus GZ60
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
AUTHORS Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L.,
Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., Butt, K.M., Wong, K.L.,
Chan, K.W., Lim, W., Shortridge, K.F., Yuen, K.Y., Peiris, J.S.M. and
Poon, L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 22913660
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 8439;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACCC 16
| | | | | | | | | | | | | | | |
DB 2420 ACCCAAGGTTTACCC 2435
```

```
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACCC 16
| | | | | | | | | | | | | | | |
DB 7115 ACCCAAGGTTTACCC 7130

RESULT 11
AY304487
LOCUS SARS coronavirus SZ13, partial genome.
DEFINITION
ACCESSION AY304487
VERSION AY304487.1 GI:34482138
KEYWORDS
SOURCE SARS coronavirus SZ13
ORGANISM SARS coronavirus SZ13
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
AUTHORS Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L.,
Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., Butt, K.M., Wong, K.L.,
Chan, K.W., Lim, W., Shortridge, K.F., Yuen, K.Y., Peiris, J.S.M. and
Poon, L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 22913660
PubMed 12958366
REFERENCE
AUTHORS Guan, Y. and Zheng, B.J.
Direct Submission
TITLE Submitted (26-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
source
1..8581
/organism="SARS coronavirus SZ13"
/mol_type="genomic RNA"
/isolate="SZ13"
/db_xref="taxon:231514"
/country="Hong Kong"
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 8581;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACCC 16
| | | | | | | | | | | | | | | |
DB 7109 ACCCAAGGTTTACCC 7124

RESULT 12
AY304491
LOCUS SARS coronavirus GZ60, partial genome.
DEFINITION
ACCESSION AY304491
VERSION AY304491.1 GI:34482142
KEYWORDS
SOURCE SARS coronavirus GZ60
ORGANISM SARS coronavirus GZ60
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
AUTHORS Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L.,
Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., Butt, K.M., Wong, K.L.,
Chan, K.W., Lim, W., Shortridge, K.F., Yuen, K.Y., Peiris, J.S.M. and
Poon, L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
MEDLINE 22913660
```

```

PUBMED 12958366
REFERENCE 2 (bases 1 to 11006)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..11006
                /organism="SARS coronavirus GZ60"
                /mol_type="genomic RNA"
                /isolate="GZ60"
                /db_xref="taxon:231518"
                /country="Hong Kong"
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 11006;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACCCAAGGTTTACCC 16
    |||||
Db 9516 ACCCAAGGTTTACCC 9531

RESULT 13
AY304493
LOCUS SARS coronavirus HKU-65806 11010 bp RNA linear VRL 05-NOV-2003
DEFINITION SARS coronavirus HKU-65806, partial genome.
ACCESSION AY304493
VERSION AY304493.1 GI:34482144
KEYWORDS
SOURCE SARS coronavirus HKU-65806
ORGANISM SARS coronavirus HKU-65806
    Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
    Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 11010)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
PUBMED 12958366
REFERENCE 2 (bases 1 to 11010)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..11010
                /organism="SARS coronavirus HKU-65806"
                /mol_type="genomic RNA"
                /isolate="HKU-65806"
                /db_xref="taxon:231520"
                /country="Hong Kong"
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 11010;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACCCAAGGTTTACCC 16
    |||||
Db 9520 ACCCAAGGTTTACCC 9535

RESULT 14
AY304494
LOCUS SARS coronavirus HKU-66078 13471 bp RNA linear VRL 05-NOV-2003
DEFINITION SARS coronavirus HKU-66078, partial genome.
ACCESSION AY304494
VERSION AY304494.1 GI:34482145
KEYWORDS
SOURCE SARS coronavirus HKU-66078
ORGANISM SARS coronavirus HKU-66078
    Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
    Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 13471)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
PUBMED 12958366
REFERENCE 2 (bases 1 to 13471)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..13471
                /organism="SARS coronavirus GZ43"
                /mol_type="genomic RNA"

```

```

DEFINITION SARS coronavirus HKU-66078, partial genome.
ACCESSION AY304494
VERSION AY304494.1 GI:34482145
KEYWORDS
SOURCE SARS coronavirus HKU-66078
ORGANISM SARS coronavirus HKU-66078
    Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
    Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 11010)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
PUBMED 12958366
REFERENCE 2 (bases 1 to 11010)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..11010
                /organism="SARS coronavirus HKU-66078"
                /mol_type="genomic RNA"
                /isolate="HKU-66078"
                /db_xref="taxon:231521"
                /country="Hong Kong"
ORIGIN
Query Match 100.0%; Score 16; DB 14; Length 11010;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACCCAAGGTTTACCC 16
    |||||
Db 9520 ACCCAAGGTTTACCC 9535

RESULT 15
AY304490
LOCUS SARS coronavirus GZ43 13471 bp RNA linear VRL 05-NOV-2003
DEFINITION SARS coronavirus GZ43, partial genome.
ACCESSION AY304490
VERSION AY304490.1 GI:34482141
KEYWORDS
SOURCE SARS coronavirus GZ43
ORGANISM SARS coronavirus GZ43
    Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
    Coronaviridae; Coronavirus.
REFERENCE 1 (bases 1 to 13471)
AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
Poon,L.L.M.
TITLE Isolation and characterization of viruses related to the SARS
coronavirus from animals in southern China
JOURNAL Science 302 (5643), 276-278 (2003)
PUBMED 12958366
REFERENCE 2 (bases 1 to 13471)
AUTHORS Guan,Y. and Zheng,B.J.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
University Pathology Building, Queen Mary Hospital, Pokfulam Road,
Hong Kong, China
FEATURES
    source
        Location/Qualifiers
            1..13471
                /organism="SARS coronavirus GZ43"
                /mol_type="genomic RNA"

```

/isolate="G243"  
/db\_xref="taxon:231517"  
/country="Hong Kong"

## ORIGIN

Query Match 100.0%; Score 16; DB 14; Length 13471;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCCCAAGGTTTACCC 16  
|||  
Db 11981 ACCCCAAGGTTTACCC 11996

Search completed: May 16, 2005, 02:00:25  
Job time : 382.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 17:17:00 ; Search time 205.25 Seconds  
(without alignments)  
461.466 Million cell updates/sec

Title: US-10-808-187A-2476

Perfect score: 16

Sequence: 1 accccaaggtttacc 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N Geneseq\_16Dec04:\*
- 1: Geneseqn1980s:\*
  - 2: Geneseqn1990s:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002as:\*
  - 7: Geneseqn2002bs:\*
  - 8: Geneseqn2003as:\*
  - 9: Geneseqn2003bs:\*
  - 10: Geneseqn2003cs:\*
  - 11: Geneseqn2003ds:\*
  - 12: Geneseqn2004as:\*
  - 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	100.0	29751	12	ADJ39000 SARS coro
2	15	93.8	3319	12	Adq22117 Human sof
3	14.4	90.0	139	3	Aaf11093 Fusarium
4	14.4	90.0	257	6	Abk76701 Bacillus
5	14.4	90.0	311	12	Adj39216 Plant cdn
6	14.4	90.0	315	4	Aai80700 Human pol
7	14.4	90.0	446	11	Adi98231 Human tum
8	14.4	90.0	473	9	Achi8176 Human adu
9	14.4	90.0	504	10	Ades5926 Rat gene
10	14.4	90.0	504	10	Ad445738 Rat gene
11	14.4	90.0	509	5	Abv54268 Human pro
12	14.4	90.0	629	11	Acn87348 Breast ca
13	14.4	90.0	663	8	Acad3754 Prokaryot
14	14.4	90.0	1161	8	ACA30490
15	14.4	90.0	1239	13	Adt45075 Bacterial
16	14.4	90.0	1497	10	Adb58261 Toxicity-
17	14.4	90.0	1497	10	Adb52793 Primary r
18	14.4	90.0	1839	12	Ado43431 Lutzomyia
19	14.4	90.0	3272	13	Ada49398 Bacterial
20	14.4	90.0	6777	4	Aac81907 P. pastor

C	21	14.4	90.0	7417	4	AAK70374	Aak70374 Human imm
	22	14.4	90.0	110000	11	ADM27081_00	Adm27081 Hyperther
	23	14.4	90.0	125910	3	AAC64370	Aac64370 Human KCN
	24	14.4	90.0	169739	6	ABQ88186	Abq88186 Human ost
C	25	14.4	90.0	334462	10	ADC24763	Adc24763 Human wil
	26	14	87.5	25	6	ABN86933	Abn86933 Human NOV
C	27	14	87.5	339	3	AAA39995	Aaa39995 Rice dise
	28	14	87.5	539	4	AAL12583	Aal12583 Human bre
C	29	14	87.5	926	11	ACN85966	Acn85966 Breast ca
	30	14	87.5	996	13	ADT47036	Adt47036 Bacterial
	31	14	87.5	1113	6	ABN86915	Abn86915 Human NOV
	32	14	87.5	1444	4	AAH44073	Aah44073 Oryza sat
	33	14	87.5	1931	13	ADS54440	Ads54440 Bacterial
	34	14	87.5	2541	8	ACC00410	Acc00410 Human cel
	35	14	87.5	2604	12	ADQ23530	Adq23530 Human sof
C	36	14	87.5	5101	4	AAK81960	Aak81960 Human imm
	37	14	87.5	5107	4	AAK81959	Aak81959 Human imm
C	38	14	87.5	5118	4	AAK81958	Aak81958 Human imm
	39	14	87.5	6292	8	ACC00270	Acc00270 Human col
	40	14	87.5	6352	8	ACC00269	Acc00269 Human col
	41	14	87.5	99046	13	ABD33291	Abd33291 Human can
	42	13.4	83.8	32	6	ABA05933	Abao5933 Human ubi
C	43	13.4	83.8	37	6	AAD46694	Adad4694 Human DBH
	44	13.4	83.8	65	6	ABN27613	Abn27613 Rat splic
C	45	13.4	83.8	272	3	AAC70965	Aac70965 Single nu

ALIGNMENTS

RESULT 1

ADJ39000

ID ADJ39000 standard; DNA; 29751 BP.

XX AC ADJ39000;

XX DT 06-MAY-2004 (first entry)

XX DE SARS coronavirus nucleotide sequence.

XX KW small interfering RNA; siRNA; modified ribonucleotide;  
viral replication inhibition; hepatitis C virus; HCV; hepatitis C;  
antiinflammatory; hepatotropic; virucide; hepatitis A virus;  
KW hepatitis D virus; hepatitis E virus; Ebola virus; influenza virus;  
KW rotavirus; reovirus; retrovirus; poliovirus; human papilloma virus;  
KW metapneumoniavirus; coronavirus; viral infection; gene; ds.

XX OS SARS coronavirus.

XX PN WO2004011647-A1.

XX PD 05-FEB-2004.

XX PF 25-JUL-2003; 2003WO-US023104.

XX PR 26-JUL-2002; 2002US-0398605P.

XX PA (CHIR ) CHIRON CORP.

XX PI Han J, Seo MY, Houghton M;

XX DR WPI; 2004-143862/14.

XX PT New RNase resistant small interfering RNA, useful for treating viral  
infections, e.g., hepatitis C, influenza virus or coronavirus infection.  
XX Example 10; Fig 3; 74pp; English.

XX CC The present invention describes a small interfering RNA (siRNA) which  
comprises a modified ribonucleotide, where the siRNA is resistant to  
RNase and retains the ability to inhibit viral replication. Also  
described: (1) inactivating a virus in a patient; (2) making a modified  
siRNA that targets a nucleic acid sequence in a virus; (3) a double-

CC stranded RNA molecule of 10-30 nucleotides that inhibits replication of  
 CC hepatitis C virus (HCV); (4) inducing targeted RNA interference toward  
 CC HCV in hepatic cells; (5) inhibiting replication of HCV; (6) a vector  
 CC comprising a DNA segment encoding the RNA molecule; (7) a host cell  
 CC comprising the vector of (6); (8) inhibiting replication of HCV in cells  
 CC carrying HCV; (9) treating hepatitis C in a subject; (10) a modified  
 CC siRNA molecule comprising a double-stranded RNA molecule of 10-30  
 CC nucleotides in length, which mediates RNA interference toward a target  
 CC agent or virus and is linked to at least one receptor-binding ligand; and  
 CC (11) inducing targeted RNA interference in a patient. The modified siRNA  
 CC molecules have antiinflammatory, hepatotropic and virucide activities.  
 CC The modified RNA molecules are useful for inactivating virus in mammalian  
 CC cells. The siRNAs are useful for treating hepatitis C virus, hepatitis A  
 CC virus, hepatitis D virus, hepatitis E virus, Ebola virus, influenza  
 CC virus, rotavirus, reovirus, coronavirus, poliovirus, human papilloma  
 CC virus, metapneumovirus or coronavirus infections. The methods of the  
 CC invention can be used to correct or compensate for cellular physiological  
 CC abnormalities involved in conferring susceptibility to viral infections  
 CC in patients and/or alleviate symptoms of a viral infection in patients.  
 CC The present sequence represents the SARS coronavirus nucleotide sequence,  
 CC which is used in an example from the present invention.

XX  
 SQ Sequence 29751 BP; 8478 A; 5941 C; 6188 G; 9144 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 16; DB 12; Length 29751;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACCC 16  
 |||||  
 Db 28245 ACCCAAGGTTTACCC 28260

RESULT 2  
 ADQ22117  
 ID ADQ22117 standard; DNA; 3319 BP.  
 XX AC ADQ22117;  
 XX  
 DT 26-AUG-2004 (first entry)  
 DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 4937.  
 XX

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;  
 KW ds.

XX Homo sapiens.  
 XX WO2004048938-A2.  
 XX  
 PD 10-JUN-2004.

XX 26-NOV-2003; 2003WO-US038193.  
 XX  
 PR 26-NOV-2002; 2002US-0429739P.  
 XX  
 PA (PROT-) PROTEIN DESIGN LABS INC.

XX Aziz N, Ginsburg WM, Zlotnik A;  
 XX WPI; 2004-441208/41.  
 XX

PT Early detection of soft tissue sarcoma comprises determining expression  
 PT of a gene in a first soft tissue sample and a normal soft tissue sample  
 PT and comparing the gene expression, also useful in treating soft tissue  
 PT sarcoma.

XX Example 2; SEQ ID NO 4937; 210pp; English.  
 PS  
 XX The invention relates to a novel method for detecting soft tissue sarcoma  
 CC which comprises obtaining a first soft tissue sample from an individual  
 CC and a normal soft tissue sample from the same or different individual,  
 CC determining the expression of a gene in both samples and comparing the

CC expression of the gene in both soft tissue samples, where a higher level  
 CC of protein expression in the first soft tissue sample indicates the  
 CC presence of soft tissue sarcoma. The method of the invention has  
 CC cytostatic applications and may be useful for detecting soft tissue  
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
 CC acid sequences may be useful in diagnostic and screening applications.  
 CC The current sequence is that of a human soft tissue sarcoma-upregulated  
 CC DNA of the invention. The current sequence is not shown within the  
 CC specification per se but was submitted in CD format by the inventor.

XX  
 SQ Sequence 3319 BP; 770 A; 828 C; 890 G; 820 T; 0 U; 11 Other;  
 Query Match 93.8%; Score 15; DB 12; Length 3319;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCCAAGGTTTACCC 16  
 |||||  
 Db 3166 CCCCAAGGTTTACCC 3180

RESULT 3  
 AAF11093/C  
 ID AAF11093 standard; cDNA; 199 BP.  
 XX  
 AC AAF11093;  
 XX  
 DT 13-MAR-2001 (first entry)  
 DE Fusarium venenatum EST SEQ ID NO:3616.

XX Multiple gene expression; filamentous fungal cell; EST;  
 KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
 KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
 KW culture condition; environmental stress; spore morphogenesis;  
 KW metabolic pathway engineering; catabolic pathway engineering; ss.

XX Fusarium venenatum.

XX WO200056762-A2.

XX 28-SEP-2000.

XX 22-MAR-2000; 2000WO-US007781.

XX 22-MAR-1999; 99US-00273623.

XX (NOVO ) NOVO NORDISK BIOTECH INC.  
 XX (NOVO ) NOVO NORDISK AS.

XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;

XX WPI; 2000-594572/56.

XX Monitoring differential expression of genes in filamentous fungal cells  
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
 PT substrate of expressed sequence tags.

XX Claim 86; Page 1645; 3161pp; English.

XX The present invention describes a method for monitoring differential  
 CC expression of genes in a first filamentous fungal (FF) cell relative to  
 CC expression of the same genes in one or more second filamentous fungal  
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
 CC are used in the methods for monitoring differential expression of genes  
 CC in a first filamentous fungal (FF) cell relative to expression of the  
 CC same genes in one or more second filamentous fungal cells. Monitoring the  
 CC global expression of genes from FF cells allows the production potential  
 CC of the microorganisms to be improved. New genes may be discovered,  
 CC possible functions of unknown open reading frames can be identified and  
 CC gene copy number variation and stability can be monitored. The expression  
 CC of genes can be used to study how FF cells adapt to changes in culture



CC conditions, environmental stress, spore morphogenesis, recombination,  
 CC metabolic or catabolic pathway engineering. Using ESTs provides several  
 CC advantages over genomic or random cDNA clones including elimination of  
 CC redundancy as one spot on an array equals one gene or open reading frame,  
 CC and organisation of the microarrays based on function of the gene  
 CC products to facilitate analysis of the results. AAF07478 to AAF11247  
 CC represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents  
 CC ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from  
 CC *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from  
 CC *Trichoderma reesei*, which are all specifically claimed in the present  
 CC invention

XX  
 SQ Sequence 199 BP; 45 A; 39 C; 44 G; 54 T; 0 U; 17 Other;

Query Match 90.0%; Score 14.4; DB 3; Length 199;  
 Best Local Similarity 93.8%; Pred. No. 5.8e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCCAGGTTTACCC 16  
 |||||  
 Db 96 ACCCCAGGTTTCC 81

## RESULT 4

ABK76701  
 ID ABK76701 standard; DNA; 257 BP.

XX  
 AC ABK76701;

XX  
 DT 13-AUG-2002 (first entry)

XX  
 DE Bacillus licheniformis genomic sequence tag (GST) #3992.

XX  
 KW Differential gene expression; genomic sequenced tag; GST;  
 KW altered culture condition; environmental stress;  
 KW physiological provocation; ds.

XX  
 XX Bacillus licheniformis.

XX  
 WO200229113-A2.

XX  
 PD 11-APR-2002.

XX  
 PF 05-OCT-2001; 2001WO-US031437.

XX  
 PR 06-OCT-2000; 2000US-00680598.

XX  
 PR 27-MAR-2001; 2001US-0279526P.

XX  
 PA (NOVO ) NOVOZYMES BIOTECH INC.

XX  
 PA (NOVO ) NOVOZYMES AS.

XX  
 PI Berka R, Clausen IG;

XX  
 DR WPI; 2002-416684/44.

XX  
 PT Monitoring differential expression of several genes in first Bacillus  
 PT cell relative to expression of same genes in one or more second Bacillus  
 PT cells, by using substrate containing Bacillus genomic sequenced tag  
 PT array.

XX  
 PS Claim 4; SEQ ID NO 3992; 200pp; English.

XX  
 CC The invention describes a method of monitoring differential expression of  
 CC genes in a first Bacillus cell relative to expression of the genes in  
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes  
 CC isolated from Bacillus cells to a substrate containing array of Bacillus  
 CC genomic sequenced tags (GST), examining the array, and determining  
 CC relative gene expression by an observed hybridisation reporter signal of  
 CC a spot in the array. The method is useful for measuring the expression of  
 CC genes in a first Bacillus cell relative to expression of the same genes  
 CC in one or more second Bacillus cells. The method is useful for monitoring  
 CC global expression of several genes from a Bacillus cell, discovering new  
 CC genes, identifying possible functions of unknown open reading frames and

CC monitoring gene copy number variation and stability. Monitoring changes  
 CC in expression of genes may be used to provide a representation of the way  
 CC in which Bacillus cells adapt to changes in culture conditions,  
 CC environmental stress or other physiological provocation. Extensive follow  
 CC up characterisation is unnecessary, when one spot on an array equals one  
 CC gene or one open reading frame, since sequence information is available.  
 CC This sequence represents a genomic sequence tag (GST) used in the method  
 CC of the invention. Note: the sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 257 BP; 65 A; 33 C; 91 G; 68 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 6; Length 257;  
 Best Local Similarity 93.8%; Pred. No. 5.9e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCCAGGTTTACCC 16  
 |||||  
 Db 233 ACCCCAGGTTAACCC 248

## RESULT 5

ADJ39216  
 ID ADJ39216 standard; cDNA; 311 BP.

XX  
 AC ADJ39216;

XX  
 DT 06-MAY-2004 (first entry)

XX  
 DE Plant cDNA #216.

XX  
 KW Plant; gene; ss; transcription; plant genome augmentation; cereal;  
 KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;  
 KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;  
 KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;  
 KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;  
 KW antifungal.

XX  
 OS Eukaryota.

XX  
 PN US2004016025-A1.

XX  
 PD 22-JAN-2004.

XX  
 PF 26-SEP-2002; 2002US-00260238.

XX  
 PR 26-SEP-2001; 2001US-0325277P.

XX  
 PR 26-SEP-2001; 2001US-0325448P.

XX  
 PR 04-APR-2002; 2002US-0370620P.

XX  
 PA (BUDW/) BUDWORTH P.

XX  
 PA (MOUG/) MOUGHAMER T.

XX  
 PA (BRIG/) BRIGGS S P.

XX  
 PA (COOP/) COOPER B.

XX  
 PA (GLAZ/) GLAZEBROOK J.

XX  
 PA (GOFF/) GOFF S A.

XX  
 PA (KATA/) KATAGIRI F.

XX  
 PA (KEEP/) KREPS J.

XX  
 PA (PROV/) PROVART N.

XX  
 PA (RICK/) RICHE D.

XX  
 PA (ZHUT/) ZHU T.

XX  
 PI Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;

XX  
 PI Goff SA, Katagiri F, Kreps J, Provart N, Riche D, Zhu T;

XX  
 DR WPI; 2004-190374/18.

XX  
 PT New rice promoter, useful for manipulating crop plants to alter or  
 PT improve phenotypic characteristics, e.g. produce large quantities of oil  
 PT or proteins, resistance to insecticides, virus or fungi, stress tolerance  
 PT or high nutritional value.

XX

PS Claim 68; SEQ ID NO 216; 230pp; English.

XX The invention relates to plant nucleotide sequences that direct seed-,

CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential

CC or constitutive transcription of an operatively linked nucleic acid

CC segment. The invention also relates to a method for augmenting a plant

CC genome and a method of identifying a gene, where its expression is

CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive

CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,

CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,

CC sorghum, rice or wheat. The polynucleotides and the polypeptides they

CC encode are useful for manipulating crop plants to alter or improve

CC phenotypic characteristics, to produce large quantities of oil or

CC proteins, to incur resistance to insecticides, viruses or fungi, and to

CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants

CC have a high nutritional value with reduced apical dominance or dwarfism,

CC early flowering or altered metabolic pathways. This sequence represents a

CC plant nucleic acid of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification but was obtained in

CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX

SQ Sequence 311 BP; 57 A; 106 C; 87 G; 61 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 12; Length 311;

Best Local Similarity 93.8%; Pred. No. 6e+02; Length 311;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16

|||||

Db 71 ACCCCAAGGTTGACCC 86

|||||

RESULT 6

AAI80700/c

ID AAI80700 standard; cDNA; 315 BP.

XX

AC AAI80700;

XX

XX

DT 06-NOV-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 760.

XX

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation; ss.

XX

OS Homo sapiens.

XX

XX WO200164835-A2.

PN

XX

PD 07-SEP-2001.

XX

XX

PF 26-FEB-2001; 2001WO-US004927.

XX

XX

PR 28-FEB-2000; 2000US-00515126.

PR

PR 18-MAY-2000; 2000US-00577409.

XX

XX (HYSE-) HYSEQ INC.

PA

XX

XX Tang YT, Liu C, Drmanac RT;

PI

XX

XX WPI; 2001-514838/56.

DR

DR P-PSDB; AAO00769.

XX

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing

PT and treating e.g. leukemia, inflammation and immune disorders.

PT

XX

XX Claim 1; SEQ ID NO 760; 1399pp + Sequence Listing; English.

PS

XX

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and

CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 315 BP; 87 A; 57 C; 89 G; 82 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 4; Length 315;

Best Local Similarity 93.8%; Pred. No. 6e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16

|||||

Db 213 ACCCCAAGGTTTCC 198

|||||

RESULT 7

ADL98231/c

ID ADL98231 standard; cDNA; 446 BP.

XX

AC ADL98231;

XX

DT 18-NOV-2004 (first entry)

XX

DE Human tumour marker gene #84.

XX

KW Human; tumour; taxane; TAXOL; sensitivity marker; resistance marker;

KW ovarian tissue; cancer; gene; ss.

XX

OS Homo sapiens.

XX

XX US2003166023-A1.

PN

XX

PD 04-SEP-2003.

XX

XX

PF 18-APR-2002; 2002US-00125159.

XX

XX

PR 18-APR-2001; 2001US-0284764P.

PR

PR 18-APR-2001; 2001US-0284773P.

PR

PR 31-MAY-2001; 2001US-0295031P.

XX

XX (MILL-) MILLENNIUM PHARM INC.

XX

XX Iartchouk N, Ayers MD, Brown JL;

PI

XX

XX WPI; 2003-898065/82.

DR

XX

XX

PT Determination of ability of TAXOL to reduce tumor growth in a patient,

PT useful to choose appropriate therapy and assess continued effectiveness

PT of therapy, uses measurement of expression of marker polynucleotides in

PT sample.

XX

XX

PS Claim 45; SEQ ID NO 84; 76pp; English.

XX

XX The invention relates to a method of determining whether the taxane

CC compound TAXOL can be used to reduce tumour growth, comprising detecting

CC expression of one or more 'sensitivity markers' or 'resistance markers'

CC in a sample of tumour cells. The method is useful for determining whether

CC TAXOL treatment will be effective to reduce tumour growth in individuals,

CC to enable appropriate therapy selection and to avoid ineffective

CC treatment, especially on a patient by patient basis. It is also useful

CC for determining whether TAXOL treatment should be continued in patients

CC (if treatment is continuing to be effective) or discontinued (if a tumour

CC has become resistant to the treatment). It may also be possible to

CC determine effectiveness or continued effectiveness of other therapeutic

CC agents/combinations of agents to which markers are sensitive/resistant,

CC and to identify new anti-cancer agents affecting marker expression. The

CC markers are useful for producing probes or primers that selectively  
CC hybridise to polynucleotides (e.g. probes hybridising to mRNA molecules),  
CC useful for detecting polynucleotides in samples, e.g. ovarian tissue  
CC (especially tumour) samples. The polypeptides encoded by the markers are  
CC useful for identifying compounds that selectively bind to the  
CC polypeptides, useful to detect polypeptides in samples. This sequence  
CC represents a tumour marker gene of the invention.

SQ Sequence 446 BP; 102 A; 81 C; 174 G; 89 T; 0 U; 0 Other;  
Query Match 90.0%; Score 14.4; DB 11; Length 446;  
Best Local Similarity 93.8%; Pred. No. 6.2e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCCAGGTTTACCC 16  
|||||  
Db 310 ACCCCAGGTTTCCCC 295

RESULT 8  
ACH18176  
ID ACH18176 standard; cDNA; 473 BP.  
XX AC ACH18176;  
XX DT 13-OCT-2003 (first entry)  
XX DE Human adult heart cDNA #2490.  
XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
XX KW genome mapping; biodiversity; genetic disorder.  
XX OS Homo sapiens.  
XX PN US2003073623-A1.  
XX PD 17-APR-2003.  
XX PF 30-JUL-2001; 2001US-00918995.  
XX PR 30-JUL-2001; 2001US-00918995.  
XX PA (DRMA/) DRMANAC R T.  
XX PA (LABA/) LABAT I.  
XX PA (STAC/) STACHE-CRAIN B.  
XX PA (DICK/) DICKSON M C.  
XX PA (JONE/) JONES L W.

PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.

XX Claim 1; SEQ ID NO 5388; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data

CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX SQ Sequence 473 BP; 142 A; 112 C; 114 G; 105 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 9; Length 473;  
Best Local Similarity 93.8%; Pred. No. 6.2e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCCAGGTTTACCC 16  
|||||  
Db 383 ACCCCAGGTTTACCC 398

RESULT 9  
ADE59926/c  
ID ADE59926 standard; DNA; 504 BP.

XX AC ADE59926;

XX DT 29-JAN-2004 (first entry)

XX DE Rat gene A892531, SEQ ID NO 5822.

XX KW Rat; ds; gene; pain; neuronal tissue; gene therapy;  
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;  
XX KW spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO ) GEN HOSPITAL CORP.

XX PA (FARB ) BAYER AG.

XX WPI; 2003-268312/26.

XX GENBANK; A892531.

XX New composition comprising two or more isolated polypeptides, useful for

XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat DNA (shown in Table 2 of the  
 CC specification) which is differentially expressed during pain. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 504 BP; 78 A; 121 C; 210 G; 95 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 10; Length 504;  
 Best Local Similarity 93.8%; Pred. No. 6.2e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACCC 16  
 |||||  
 Db 406 ACCCAAGGTTTACCC 391

RESULT 10  
 ADD45738/c  
 ID ADD45738 standard; DNA; 504 BP.  
 AC ADD45738;  
 XX

DT 29-JAN-2004 (first entry)

DE Rat gene AA892531, SEQ ID NO 11406.

KW Rat; ds; gene; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

XX WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; AA892531.

PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat DNA (shown in Table 2 of the  
 CC specification) which encodes one of the polypeptides of the invention  
 CC which is differentially expressed during pain. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 504 BP; 78 A; 121 C; 210 G; 95 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 10; Length 504;  
 Best Local Similarity 93.8%; Pred. No. 6.2e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACCC 16  
 |||||  
 Db 406 ACCCAAGGTTTACCC 391

RESULT 11

ABV54268/c

ID ABV54268 standard; cDNA; 509 BP.

AC ABV54268;

XX 17-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 54259.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 10495; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 509 BP; 196 A; 77 C; 81 G; 152 T; 0 U; 3 Other;

Query Match 90.0%; Score 14.4; DB 5; Length 509;  
 Best Local Similarity 93.8%; Pred. No. 6.2e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCCAGGTTTACCC 16  
 |||||  
 Db 289 ACCCCAGGTTTCC 274

## RESULT 12

ACN87348/c  
 ID ACN87348 standard; DNA; 629 BP.

XX AC ACN87348;

XX 02-DEC-2004 (first entry)

XX Breast cancer related marker, seq id 8498.

XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX Homo sapiens.

XX US2003099974-A1.

XX 29-MAY-2003.

XX 18-JUL-2002; 2002US-00198846.

XX 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2003-787014/74.

XX Novel isolated polypeptide associated with breast cancer, useful for  
 PT detecting presence of polypeptide in sample, as a marker for breast  
 PT cancer.

XX Disclosure; SEQ ID NO 8498; 36pp; English.

PS The invention relates to an isolated polypeptide (I) associated with  
 CC breast cancer which is encoded by a nucleic acid molecule comprising a  
 CC nucleotide sequence (SI). Further disclosed is an antibody that binds to  
 CC the polypeptide of the invention. The activity of the polypeptide of the  
 CC invention may be described as cytostatic. The antibody is useful for  
 CC detecting the presence of (I) in a sample. Nucleic acid molecules of the  
 CC invention are useful in the detection of breast tumours. (I) is useful as  
 CC a marker for breast cancer and in breast cancer therapy. Sequences given  
 CC in records ACN78851-ACN92934 represent nucleic acid markers associated  
 CC with breast cancer. Note: The sequence listing does not form part of the  
 CC specification but may be obtained in electronic format from the USPTO web  
 CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974

XX Sequence 629 BP; 161 A; 161 C; 159 G; 136 T; 0 U; 12 Other;

Query Match 90.0%; Score 14.4; DB 11; Length 629;  
 Best Local Similarity 93.8%; Pred. No. 6.3e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCCAGGTTTACCC 16  
 |||||  
 Db 611 ACCCCAGGTTTCC 596

## RESULT 13

ACA43754/c  
 ID ACA43754 standard; DNA; 663 BP.

XX AC ACA43754;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #25411.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
 XX drug design; gene.

XX Pseudomonas putida.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 08-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-0299926/02.

XX P-PSDB; ABU39884.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 31624; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 663 BP; 140 A; 218 C; 193 G; 112 T; 0 U; 0 Other;  
  
Query Match 90.0%; Score 14.4; DB 8; Length 663;  
Best Local Similarity 93.8%; Pred. NO. 6.3e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ACCCAAGGTTTACCC 16  
||| ||||| |||||  
Db 548 ACCTCAAGGTTTACCC 533  
  
RESULT 14  
ID ACA30490 standard; DNA; 1161 BP.  
XX  
AC ACA30490;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Prokaryotic essential gene #12147.  
XX  
KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
XX  
OS Campylobacter jejuni.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
XX  
WPI; 2003-029926/02.  
DR P-PSDB; ABU26620.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 18360; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
  
to which each of the strains is present in a culture or collection of  
strains; or (13) identifying the target of a compound that inhibits the  
proliferation of an organism. The antisense nucleic acids are useful for  
identifying proteins or screening for homologous nucleic acids required  
for cellular proliferation to isolate candidate molecules for rational  
drug discovery programs, or for screening homologous nucleic acids  
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
*K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
prokaryotic essential genes. Note: the sequence data for this patent did  
not form part of the printed specification, but was obtained in  
electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences  
  
SQ Sequence 1161 BP; 426 A; 174 C; 200 G; 361 T; 0 U; 0 Other;  
  
Query Match 90.0%; Score 14.4; DB 8; Length 1161;  
Best Local Similarity 93.8%; Pred. NO. 6.6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ACCCAAGGTTTACCC 16  
||| ||||| |||||  
Db 519 ACCCAAGGTTTACCC 504  
  
RESULT 15  
ADT45075/c  
ID ADT45075 standard; cDNA; 1239 BP.  
XX  
AC ADT45075;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polynucleotide #19826.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polynucleotide; gene; ss.  
XX  
OS Bacteria.  
XX  
PN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
XX (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 43513; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polynucleotide used in  
CC the scope of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX

SQ Sequence 1239 BP; 342 A; 278 C; 318 G; 301 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 13; Length 1239;  
Best Local Similarity 93.8%; Pred. No. 6.7e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACCCCAAGGTTTACCC 16  
| | | | | | | | | | | | | | | |  
Db 1191 ACCCCAAGGTTTACCC 1176

Search completed: May 16, 2005, 01:09:40  
Job time : 210.25 secs

**"THIS PAGE BLANK (USPTO)"**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:11:05 ; Search time 59.875 Seconds  
(without alignments)  
437.251 Million cell updates/sec

Title: US-10-808-187A-2476  
Perfect score: 16  
Sequence: 1 accccaaggtttacc 16  
Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14.4	90.0	39154	4	US-09-949-016-12384
C 2	14.4	90.0	39154	4	US-09-949-016-12801
C 3	14.4	90.0	39443	4	US-09-949-016-14326
C 4	14.4	90.0	39443	4	US-09-949-016-14327
C 5	14	87.5	26896	4	US-09-949-016-16800
C 6	13.4	83.8	373	4	US-09-270-767-7591
C 7	13.4	83.8	373	4	US-09-270-767-22873
C 8	13.4	83.8	601	4	US-09-949-016-173609
C 9	13.4	83.8	624	3	US-09-328-111-146
C 10	13.4	83.8	735	4	US-09-543-681A-423
C 11	13.4	83.8	1206	4	US-09-540-236-1241
C 12	13.4	83.8	1270	4	US-09-799-451-81
C 13	13.4	83.8	1512	4	US-09-408-020-65
C 14	13.4	83.8	1614	4	US-09-602-777A-409
C 15	13.4	83.8	1679	2	US-08-365-486A-11
C 16	13.4	83.8	1679	3	US-08-880-342-11
C 17	13.4	83.8	1991	1	US-08-295-814E-9
C 18	13.4	83.8	1991	3	US-09-343-361-9
C 19	13.4	83.8	1991	5	PCT-US93-01959-9
C 20	13.4	83.8	2025	4	US-09-543-681A-1819
C 21	13.4	83.8	2448	4	US-09-328-352-158
C 22	13.4	83.8	2814	4	US-09-583-110-505
C 23	13.4	83.8	2928	4	US-09-107-433-429
C 24	13.4	83.8	4327	3	US-08-961-527-117
C 25	13.4	83.8	4500	2	US-08-743-637B-35
C 26	13.4	83.8	4500	3	US-08-526-840B-35
C 27	13.4	83.8	5153	5	PCT-US95-04910-8

C 28	13.4	83.8	8224	2	US-09-010-398-14	Sequence 14, Appl
C 29	13.4	83.8	8224	3	US-09-366-260-14	Sequence 14, Appl
C 30	13.4	83.8	9008	4	US-09-949-016-12576	Sequence 12576, A
C 31	13.4	83.8	9009	4	US-09-949-016-14036	Sequence 14036, A
C 32	13.4	83.8	21000	4	US-09-975-123-11	Sequence 11, Appl
C 33	13.4	83.8	22908	4	US-09-949-016-17255	Sequence 17255, A
C 34	13.4	83.8	30847	4	US-09-949-016-16657	Sequence 16657, A
C 35	13.4	83.8	32998	4	US-09-408-020-1	Sequence 1, Appl
C 36	13.4	83.8	33248	4	US-09-596-002-24	Sequence 24, Appl
C 37	13.4	83.8	60276	4	US-09-949-016-15004	Sequence 15004, A
C 38	13.4	83.8	60338	4	US-09-949-016-15694	Sequence 15694, A
C 39	13.4	83.8	183770	4	US-09-949-016-15494	Sequence 15494, A
C 40	13.4	83.8	325034	4	US-09-949-016-14957	Sequence 14957, A
C 41	13.4	83.8	389504	4	US-09-949-016-11774	Sequence 11774, A
C 42	13.4	83.8	767677	4	US-09-949-016-12147	Sequence 12147, A
C 43	13.4	83.8	767677	4	US-09-949-016-17361	Sequence 17361, A
C 44	13	81.2	419	4	US-09-621-976-14404	Sequence 14404, A
C 45	13	81.2	497	4	US-09-621-976-10528	Sequence 10528, A

ALIGNMENTS

RESULT 1  
US-09-949-016-12384/c  
; Sequence 12384, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12384  
; LENGTH: 39154  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(39154)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12384  
  
Query Match 90.0%; Score 14.4; DB 4; Length 39154;  
Best Local Similarity 93.8%; Pred. No. 2.9e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ACCCCAAGGTTTACCC 16  
Db 31781 ACCCAAGGTTTACCC 31766  
  
RESULT 2  
US-09-949-016-12801/c  
; Sequence 12801, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12801
; LENGTH: 39154
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39154)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12801
```

```
Query Match          90.0%; Score 14.4; DB 4; Length 39154;
Best Local Similarity 93.8%; Pred. No. 2.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 ACCCCAAGGTTTACCC 16
      ||||| ||||| |||||
Db      31781 ACCCCACGGTTTACCC 31766
```

## RESULT 3

```
US-09-949-016-14326/c
; Sequence 14326, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14326
; LENGTH: 39443
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39443)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14326
```

```
Query Match          90.0%; Score 14.4; DB 4; Length 39443;
Best Local Similarity 93.8%; Pred. No. 2.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 ACCCCAAGGTTTACCC 16
      ||||| ||||| |||||
Db      31781 ACCCCACGGTTTACCC 31766
```

## RESULT 4

```
US-09-949-016-14327/c
; Sequence 14327, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14327
; LENGTH: 39443
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39443)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14327
```

```
Query Match          90.0%; Score 14.4; DB 4; Length 39443;
Best Local Similarity 93.8%; Pred. No. 2.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 ACCCCAAGGTTTACCC 16
      ||||| ||||| |||||
Db      31781 ACCCCACGGTTTACCC 31766
```

## RESULT 5

```
US-09-949-016-16800
; Sequence 16800, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16800
; LENGTH: 26896
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16800
```

```
Query Match          87.5%; Score 14; DB 4; Length 26896;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ACCCCAAGGTTTAC 14
      ||||| ||||| |||||
Db      10157 ACCCCAAGGTTTAC 10170
```

## RESULT 6

```
US-09-270-767-7591/c
; Sequence 7591, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
```

```
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7591  
; LENGTH: 373  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-7591

Query Match 83.8%; Score 13.4; DB 4; Length 373;  
Best Local Similarity 93.3%; Pred. No. 6.7e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCAAGGTTTACC 16  
| | | | | | | | | | | | | | | |  
DB 46 CCGCAAGGTTTACC 32

RESULT 7

US-09-270-767-22873/c  
; Sequence 22873, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22873  
; LENGTH: 373  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-22873

Query Match 83.8%; Score 13.4; DB 4; Length 373;  
Best Local Similarity 93.3%; Pred. No. 6.7e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCAAGGTTTACC 16  
| | | | | | | | | | | | | | | |  
DB 46 CCGCAAGGTTTACC 32

RESULT 8

US-09-949-016-173609  
; Sequence 173609, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 173609  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-173609

Query Match 83.8%; Score 13.4; DB 4; Length 601;  
Best Local Similarity 93.3%; Pred. No. 7e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACC 15  
| | | | | | | | | | | | | | | |  
DB 583 AACCAAGGTTTACC 597

RESULT 9

US-09-328-111-146/c  
; Sequence 146, Application US/09328111  
; Patent No. 6262333  
; GENERAL INFORMATION:  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Steinmann, Kathleen E.  
; APPLICANT: Ascle, Jon H.  
; APPLICANT: Burgess, Christopher C.  
; APPLICANT: Bushnell, Steven E.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Catino, Theodore J.  
; APPLICANT: Derti, Adnan  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Lewis, Marcia E.  
; APPLICANT: Monahan, John E.  
; APPLICANT: Schlegel, Robert  
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
; FILE REFERENCE: CCD-257 (US)  
; CURRENT APPLICATION NUMBER: US/09/328,111  
; CURRENT FILING DATE: 1999-06-08  
; EARLIER APPLICATION NUMBER: US 60/088,801  
; EARLIER FILING DATE: 1998-06-10  
; NUMBER OF SEQ ID NOS: 850  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 146  
; LENGTH: 624  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(624)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-328-111-146

Query Match 83.8%; Score 13.4; DB 3; Length 624;  
Best Local Similarity 93.3%; Pred. No. 7e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCAAGGTTTACC 16  
| | | | | | | | | | | | | | | |  
DB 624 CCCCAAGGTTTACC 610

RESULT 10

US-09-543-681A-423/c  
; Sequence 423, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 423  
; LENGTH: 735  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-423

Query Match 83.8%; Score 13.4; DB 4; Length 735;  
Best Local Similarity 93.3%; Pred. No. 7.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCAAGGTTTACCC 16  
|||||  
Db 176 CCCCAAGGTTTAGCC 162

RESULT 11  
US-09-540-236-1241  
; Sequence 1241, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 1241  
; LENGTH: 1206  
; TYPE: DNA  
; ORGANISM: M.cattarrhalis  
US-09-540-236-1241

Query Match 83.8%; Score 13.4; DB 4; Length 1206;  
Best Local Similarity 93.3%; Pred. No. 7.4e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCAAGGTTTACCC 16  
|||||  
Db 983 CCCCAAGGTTTACCC 997

RESULT 12  
US-09-799-451-81  
; Sequence 81, Application US/09799451  
; Patent No. 6783969

; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Ma, Yungqing  
; APPLICANT: Yamazaki, Victoria  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Ghosh, Reena  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 803  
; CURRENT APPLICATION NUMBER: US/09/799,451  
; CURRENT FILING DATE: 2001-03-05  
; NUMBER OF SEQ ID NOS: 948  
; SOFTWARE: pt\_FL\_genes Version 2.0  
; SEQ ID NO 81  
; LENGTH: 1270  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (63)..(1124)  
US-09-799-451-81

Query Match 83.8%; Score 13.4; DB 4; Length 1270;

Best Local Similarity 93.3%; Pred. No. 7.5e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCAAGGTTTACCC 16  
|||||  
Db 841 CCCCAAGGTTTACCC 855

RESULT 13  
US-09-408-020-65/c  
; Sequence 65, Application US/09408020  
; Patent No. 6632937  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCORP.002A  
; CURRENT APPLICATION NUMBER: US/09/408,020  
; CURRENT FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: 60/102,294  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 65  
; LENGTH: 1512  
; TYPE: DNA  
; ORGANISM: Cenarchaeum symbiosum  
; NAME/KEY: CDS  
; LOCATION: (1)....(1512)  
US-09-408-020-65

Query Match 83.8%; Score 13.4; DB 4; Length 1512;  
Best Local Similarity 93.3%; Pred. No. 7.6e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACC 15  
|||  
Db 1268 ACCCAAGGTTTACC 1254

RESULT 14  
US-09-602-777A-409  
; Sequence 409, Application US/09602777A  
; Patent No. 6831165  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schroder, Hartwig  
; APPLICANT: Zeider, Oskar  
; APPLICANT: Haberhauer, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION  
; FILE REFERENCE: BGI-128CP  
; CURRENT APPLICATION NUMBER: US/09/602,777A  
; CURRENT FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/141031  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: DE 19931636.8  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19932125.6  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932126.4  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932127.2  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932128.0  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932129.9  
; PRIOR FILING DATE: 1999-07-19  
; PRIOR APPLICATION NUMBER: DE 19932226.0  
; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932920.6  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932922.2  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932924.9  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932928.1  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932930.3  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932933.8  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932935.4  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932973.7  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19933002.6  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19933003.4  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19933005.0  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19933006.9  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19941378.9  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: DE 19941379.7  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: DE 19941390.8  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: DE 19941391.6  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: DE 19942088.2  
; PRIOR FILING DATE: 1999-09-03  
; NUMBER OF SEQ ID NOS: 442  
; SEQ ID NO 409  
; LENGTH: 1614  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (101)..(1591)  
; OTHER INFORMATION: RXN00641  
US-09-602-777A-409

Query Match 83.8%; Score 13.4; DB 4; Length 1614;  
Best Local Similarity 93.3%; Pred. No. 7.6e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACC 15  
Db 1041 ACCCAAGTGTTCAC 1055

## RESULT 15

US-08-365-486A-11/c  
; Sequence 11, Application US/08365486A  
; Patent No. 5834306  
; GENERAL INFORMATION:  
; APPLICANT: Webster, Keith A.  
; APPLICANT: Bishopric, Nanette H.  
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated  
; TITLE OF INVENTION: Therapeutic Constructs  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25.  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/365,486A  
; FILING DATE: 23-DEC-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: 38,615  
; REFERENCE/DOCKET NUMBER: 8255-0018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1679 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Mouse alpha MHC promoter fragment  
US-08-365-486A-11

Query Match 83.8%; Score 13.4; DB 2; Length 1679;  
Best Local Similarity 93.3%; Pred. No. 7.6e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCCAAGGTTTACC 16  
Db 447 CCCCAAGTGTTCAC 433

Search completed: May 16, 2005, 06:11:47  
Job time : 62.875 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 02:00:36 ; Search time 758.5 Seconds  
(without alignments)  
129.102 Million cell updates/sec

Title: US-10-808-187A-2476

Perfect score: 16

Sequence: 1 accccaaggtttacc 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	100.0	16	19	US-10-808-187-2476
2	16	100.0	1620	19	US-10-699-936-16
3	16	100.0	2304	19	US-10-699-936-7
4	16	100.0	2304	19	US-10-699-936-11
5	16	100.0	24774	19	US-10-889-447-3
6	16	100.0	28920	19	US-10-889-447-5
7	16	100.0	28920	19	US-10-889-447-6
8	16	100.0	29291	19	US-10-889-447-4
9	16	100.0	29430	19	US-10-889-447-7
10	16	100.0	29727	18	US-10-839-729-15
11	16	100.0	29727	18	US-10-827-757-1

12	16	100.0	29727	19	US-10-889-447-8
13	16	100.0	29727	19	US-10-699-936-1
14	16	100.0	29736	18	US-10-839-729-17
15	16	100.0	29736	19	US-10-889-447-9
16	16	100.0	29736	19	US-10-699-936-3
17	16	100.0	29742	18	US-10-839-729-16
18	16	100.0	29742	19	US-10-808-187-15
19	16	100.0	29742	19	US-10-808-187-16
20	16	100.0	29742	19	US-10-808-187-240
21	16	100.0	29742	19	US-10-808-187-737
22	16	100.0	29742	19	US-10-808-187-1108
C 23	16	100.0	29742	19	US-10-808-187-1590
C 24	16	100.0	29742	19	US-10-808-187-1965
25	16	100.0	29742	19	US-10-889-447-10
26	16	100.0	29751	18	US-10-839-729-14
27	16	100.0	29751	19	US-10-856-529-1
28	16	100.0	29751	19	US-10-626-879-67
29	16	100.0	29751	19	US-10-889-447-1
30	16	100.0	29751	19	US-10-889-447-2
31	16	100.0	29751	19	US-10-699-936-2
32	15	93.8	634	17	US-10-424-599-140261
33	15	93.8	3319	18	US-10-723-860-4937
C 34	14.4	90.0	157	17	US-10-424-599-40799
C 35	14.4	90.0	159	18	US-10-653-047-3616
C 36	14.4	90.0	235	17	US-10-424-599-67863
C 37	14.4	90.0	241	18	US-10-437-963-86252
38	14.4	90.0	257	9	US-09-974-300-3992
39	14.4	90.0	311	17	US-10-260-238-216
40	14.4	90.0	311	18	US-10-437-963-66704
41	14.4	90.0	345	17	US-10-424-599-89379
C 42	14.4	90.0	362	17	US-10-424-599-101240
C 43	14.4	90.0	391	18	US-10-425-115-126942
C 44	14.4	90.0	396	18	US-10-437-963-30933
C 45	14.4	90.0	398	19	US-10-950-009-798

ALIGNMENTS

RESULT 1

- US-10-808-187-2476
- Sequence 2476, Application US/10808187
- Publication No. US2005000909A1
- GENERAL INFORMATION:
- APPLICANT: PEIRIS, JOSEPH S. M.
- APPLICANT: YUEN, KWOK YUNG
- APPLICANT: POON, LIT MAN
- APPLICANT: GUAN, YI
- APPLICANT: CHAN, KWOK HUNG
- APPLICANT: NICHOLLS, JOHN
- TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE RESPIRATORY SYNDROME (SARS)
- TITLE OF INVENTION: RESPIRATORY SYNDROME (SARS)
- FILE REFERENCE: V9661.0078
- CURRENT APPLICATION NUMBER: US/10/808,187
- CURRENT FILING DATE: 2004-03-24
- PRIOR APPLICATION NUMBER: 60/457,031
- PRIOR FILING DATE: 2003-03-24
- PRIOR APPLICATION NUMBER: 60/457,730
- PRIOR FILING DATE: 2003-03-26
- PRIOR APPLICATION NUMBER: 60/459,931
- PRIOR FILING DATE: 2003-04-02
- PRIOR APPLICATION NUMBER: 60/460,357
- PRIOR FILING DATE: 2003-04-03
- PRIOR APPLICATION NUMBER: 60/461,265
- PRIOR FILING DATE: 2003-04-08
- PRIOR APPLICATION NUMBER: 60/462,805
- PRIOR FILING DATE: 2003-04-14
- PRIOR APPLICATION NUMBER: 60/468,139
- PRIOR FILING DATE: 2003-05-05
- PRIOR APPLICATION NUMBER: 60/464,886
- PRIOR FILING DATE: 2003-04-23
- PRIOR APPLICATION NUMBER: 60/471,200
- PRIOR FILING DATE: 2003-05-16

; NUMBER OF SEQ ID NOS: 2476  
; SOFTWARE: PatentIn ver. 3.2  
; SEQ ID NO 2476  
; LENGTH: 16  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Probe  
US-10-808-187-2476

Query Match 100.0%; Score 16; DB 19; Length 16;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16  
|||||  
DB 1 ACCCCAAGGTTTACCC 16

RESULT 2  
US-10-699-936-16  
; Sequence 16, Application US/10699936  
; Publication No. US20050095582A1  
; GENERAL INFORMATION:  
; APPLICANT: Gillim-Ross, Laura  
; APPLICANT: Taylor, Jill  
; APPLICANT: Scholl, David R.  
; APPLICANT: Wentworth, David E.  
; APPLICANT: Jollick, Joseph D.  
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory  
; TITLE OF INVENTION: Syndrome Coronavirus  
; FILE REFERENCE: DHI-07986  
; CURRENT APPLICATION NUMBER: US/10/699,936  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 16  
; LENGTH: 1620  
; TYPE: DNA  
; ORGANISM: SARS-CoV ZJ-HZ01  
US-10-699-936-16

Query Match 100.0%; Score 16; DB 19; Length 1620;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16  
|||||  
DB 200 ACCCCAAGGTTTACCC 215

RESULT 3  
US-10-699-936-7  
; Sequence 7, Application US/10699936  
; Publication No. US20050095582A1  
; GENERAL INFORMATION:  
; APPLICANT: Gillim-Ross, Laura  
; APPLICANT: Taylor, Jill  
; APPLICANT: Scholl, David R.  
; APPLICANT: Wentworth, David E.  
; APPLICANT: Jollick, Joseph D.  
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory  
; TITLE OF INVENTION: Syndrome Coronavirus  
; FILE REFERENCE: DHI-07986  
; CURRENT APPLICATION NUMBER: US/10/699,936  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7  
; LENGTH: 2304  
; TYPE: DNA  
; ORGANISM: SARS-CoV Shanghai LY  
US-10-699-936-7

Query Match 100.0%; Score 16; DB 19; Length 2304;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16  
|||||  
DB 880 ACCCCAAGGTTTACCC 895

RESULT 4  
US-10-699-936-11  
; Sequence 11, Application US/10699936  
; Publication No. US20050095582A1  
; GENERAL INFORMATION:  
; APPLICANT: Gillim-Ross, Laura  
; APPLICANT: Taylor, Jill  
; APPLICANT: Scholl, David R.  
; APPLICANT: Wentworth, David E.  
; APPLICANT: Jollick, Joseph D.  
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory  
; TITLE OF INVENTION: Syndrome Coronavirus  
; FILE REFERENCE: DHI-07986  
; CURRENT APPLICATION NUMBER: US/10/699,936  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 11  
; LENGTH: 2304  
; TYPE: DNA  
; ORGANISM: SARS coronavirus Shanghai LY  
US-10-699-936-11

Query Match 100.0%; Score 16; DB 19; Length 2304;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16  
|||||  
DB 880 ACCCCAAGGTTTACCC 895

RESULT 5  
US-10-889-447-3  
; Sequence 3, Application US/10889447  
; Publication No. US20050075307A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Jain, Ravi  
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
; FILE REFERENCE: RTS-0685US  
; CURRENT APPLICATION NUMBER: US/10/889,447  
; CURRENT FILING DATE: 2004-07-12  
; PRIOR APPLICATION NUMBER: 60/486,670  
; PRIOR FILING DATE: 2003-07-12  
; NUMBER OF SEQ ID NOS: 241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 24774  
; TYPE: DNA  
; ORGANISM: SARS coronavirus isolate BJ01  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION:  
; OTHER INFORMATION: n is any nucleotide  
US-10-889-447-3

Query Match 100.0%; Score 16; DB 19; Length 24774;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16  
|||||



```
Db      23358 ACCCAAGGTTTACCC 23373

RESULT 6
US-10-889-447-5
; Sequence 5, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ03
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-5
Query Match      100.0%; Score 16; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCAAGGTTTACCC 16
      |||||||
Db      27507 ACCCAAGGTTTACCC 27522

RESULT 7
US-10-889-447-6
; Sequence 6, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ04
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-6
Query Match      100.0%; Score 16; DB 19; Length 28920;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCAAGGTTTACCC 16
      |||||||
Db      27507 ACCCAAGGTTTACCC 27522

RESULT 8
US-10-889-447-4
```

```
; Sequence 4, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 29291
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ02
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-4
Query Match      100.0%; Score 16; DB 19; Length 29291;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCAAGGTTTACCC 16
      |||||||
Db      27875 ACCCAAGGTTTACCC 27890

RESULT 9
US-10-889-447-7
; Sequence 7, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 29430
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate GZ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-7
Query Match      100.0%; Score 16; DB 19; Length 29430;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCAAGGTTTACCC 16
      |||||||
Db      28010 ACCCAAGGTTTACCC 28025

RESULT 10
US-10-839-729-15
; Sequence 15, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
```

```
; TITLE OF INVENTION: AND METHODS OF USE
; FILE REFERENCE: BIOBANK.013A
; CURRENT APPLICATION NUMBER: US/10/839,729
; CURRENT FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: 60/468703
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-839-729-15

Query Match      100.0%; Score 16; DB 18; Length 29727;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCCAAGGTTTACCC 16
      |||||
Db      28245 ACCCCAAGGTTTACCC 28260

RESULT 11
US-10-827-757-1
; Sequence 1, Application US/10827757
; Publication No. US20050004071A1
; GENERAL INFORMATION:
; APPLICANT: Comper, Wayne
; TITLE OF INVENTION: Charged Polysaccharides Resistant To Lysosomal Degradation During
; TITLE OF INVENTION: Kidney Filtration And Renal Passage And Their Use To Treat Or
; FILE REFERENCE: 11213-007-999
; CURRENT APPLICATION NUMBER: US/10/827,757
; CURRENT FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 60/464,294
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS-related coronavirus (Urbani strain)
US-10-827-757-1

Query Match      100.0%; Score 16; DB 18; Length 29727;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCCAAGGTTTACCC 16
      |||||
Db      28245 ACCCCAAGGTTTACCC 28260

RESULT 12
US-10-889-447-8
; Sequence 8, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS coronavirus Urbani

Query Match      100.0%; Score 16; DB 18; Length 29736;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCCAAGGTTTACCC 16
      |||||
Db      28230 ACCCCAAGGTTTACCC 28245

US-10-889-447-8
Query Match      100.0%; Score 16; DB 19; Length 29727;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCCAAGGTTTACCC 16
      |||||
Db      28245 ACCCCAAGGTTTACCC 28260

RESULT 13
US-10-699-936-1
; Sequence 1, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS coronavirus Urbani
US-10-699-936-1

Query Match      100.0%; Score 16; DB 19; Length 29727;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCCAAGGTTTACCC 16
      |||||
Db      28245 ACCCCAAGGTTTACCC 28260

RESULT 14
US-10-839-729-17
; Sequence 17, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
; TITLE OF INVENTION: AND METHODS OF USE
; FILE REFERENCE: BIOBANK.013A
; CURRENT APPLICATION NUMBER: US/10/839,729
; CURRENT FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: 60/468703
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 29736
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-839-729-17

Query Match      100.0%; Score 16; DB 18; Length 29736;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACCCCAAGGTTTACCC 16
      |||||
Db      28230 ACCCCAAGGTTTACCC 28245
```

RESULT 15  
US-10-889-447-9  
; Sequence 9, Application US/10889447  
; Publication No. US20050075307A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Jain, Ravi  
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION  
; FILE REFERENCE: RTS-06850S  
; CURRENT APPLICATION NUMBER: US/10/889,447  
; CURRENT FILING DATE: 2004-07-12  
; PRIOR APPLICATION NUMBER: 60/486,670  
; PRIOR FILING DATE: 2003-07-12  
; NUMBER OF SEQ ID NOS: 241  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 29736  
; TYPE: DNA  
; ORGANISM: SARS coronavirus CUHK-W1  
US-10-889-447-9

Query Match 100.0%; Score 16; DB 19; Length 29736;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCAAGGTTTACCC 16  
|||||  
Db 28230 ACCCAAGGTTTACCC 28245

Search completed: May 16, 2005, 14:33:52  
Job time : 760.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 00:01:20 ; Search time 1807 Seconds  
(without alignments)  
337.038 Million cell updates/sec

Title: US-10-808-187A-2476

Perfect score: 16

Sequence: 1 accccaaggtttacc 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16	100.0	427	5	BP653652
C 2	16	100.0	514	9	CR488879
C 3	16	100.0	543	9	CR478789
C 4	16	100.0	567	9	CR312442
C 5	16	100.0	606	9	CR348467
C 6	16	100.0	759	6	CB986444
C 7	16	100.0	788	8	AZ553628
C 8	16	100.0	1017	5	BU931691
C 9	15	93.8	391	6	CD888930
C 10	15	93.8	412	6	CA706577
C 11	15	93.8	425	5	BP660567
C 12	15	93.8	480	2	BE423723
C 13	15	93.8	496	6	CD843729
C 14	15	93.8	499	5	BP641522
C 15	15	93.8	528	4	BJ538102
C 16	15	93.8	549	2	BF416296
C 17	15	93.8	555	4	BJ525536
C 18	15	93.8	567	4	BJ003177
C 19	15	93.8	572	6	CD899758
C 20	15	93.8	576	4	BJ009030
C 21	15	93.8	577	2	BF416297
C 22	15	93.8	588	4	BJ521627
C 23	15	93.8	591	6	CA604979
C 24	15	93.8	600	4	BJ520055

C 25	15	93.8	606	4	BJ532637
C 26	15	93.8	619	4	BJ014018
C 27	15	93.8	624	6	CA238978
C 28	15	93.8	633	4	BJ005175
C 29	15	93.8	634	4	BJ003002
C 30	15	93.8	644	4	BJ013066
C 31	15	93.8	646	4	BJ012095
C 32	15	93.8	646	4	BJ535173
C 33	15	93.8	647	4	BJ536816
C 34	15	93.8	650	5	BU942217
C 35	15	93.8	651	4	BJ028177
C 36	15	93.8	651	4	BJ022751
C 37	15	93.8	658	4	BJ019178
C 38	15	93.8	661	1	AI525268
C 39	15	93.8	662	4	BJ017517
C 40	15	93.8	671	4	BJ027351
C 41	15	93.8	675	5	BQ783177
C 42	15	93.8	675	6	CA456029
C 43	15	93.8	686	8	BH930361
C 44	15	93.8	692	4	BU524239
C 45	15	93.8	699	8	BZ115852

ALIGNMENTS

RESULT 1  
BP653652/c  
LOCUS BP653652 RAFL19 Arabidopsis thaliana cDNA clone RAFL19-14-M03 3',  
DEFINITION mRNA sequence.  
ACCESSION BP653652  
VERSION BP653652.1 GI:49305122  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1. (bases 1 to 427)  
AUTHORS Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M., Hayaishizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.

Functional annotation of a full-length Arabidopsis cDNA collection  
Science 296 (5565), 141-145 (2002)  
TITLE JOURNAL  
MEDLINE 21932900  
PubMed 11910074

Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: mseki@rtc.riken.go.jp  
Reversed clone; Please visit our web site  
(http://pfweb.gsc.riken.go.jp/) for further details.

FEATURES  
source  
Location/Qualifiers  
1..427  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAFL19-14-M03"  
/tissue\_type="mixture of silique and flower"  
/lab\_host="DH10B"  
/clone\_lib="RAFL19"  
/note="Site\_1: BamHI; Site\_2: SalI; Subtraction library"

ORIGIN  
Query Match 100.0%; Score 16; DB 5; Length 427;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ACCCCAAGGTTTACCC 16
Db 301 ACCCCAAGGTTTACCC 286

RESULT 2
LOCUS CR48879/c 514 bp DNA linear GSS 11-JUN-2004
DEFINITION Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago
truncatula, genomic survey sequence.
ACCESSION CR48879
VERSION CR48879.1 GI:48650455
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 514)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
FEATURES
source
Location/Qualifiers
1..514
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="Jemalong A17"
/db_xref="taxon:3880"
/clone_lib="MTH2"
/note="Vector: pBelobAC11 ; Site 1: HindIII ; Site 2:
HindIII ; Cook, D.R. and Kim, D.J-Genoscope sequence ID :
mth2-159A6FM1"

ORIGIN
Query Match 100.0%; Score 16; DB 9; Length 514;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16
Db 28 ACCCCAAGGTTTACCC 13

RESULT 3
LOCUS CR478789/c 543 bp DNA linear GSS 11-JUN-2004
DEFINITION Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago
truncatula, genomic survey sequence.
ACCESSION CR478789
VERSION CR478789.1 GI:48640365
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 543)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
FEATURES
source
Location/Qualifiers
1..543
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="Jemalong A17"
/db_xref="taxon:3880"
/clone_lib="MTH2"

```

```

/note="Vector: pBelobAC11 ; Site 1: HindIII ; Site 2:
HindIII ; Cook, D.R. and Kim, D.J-Genoscope sequence ID :
mth2-188J7FM1"

ORIGIN
Query Match 100.0%; Score 16; DB 9; Length 543;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16
Db 28 ACCCCAAGGTTTACCC 13

RESULT 4
LOCUS CR312442 567 bp DNA linear GSS 01-MAR-2004
DEFINITION Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago
truncatula, genomic survey sequence.
ACCESSION CR312442
VERSION CR312442.1 GI:44858586
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 567)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
FEATURES
source
Location/Qualifiers
1..567
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="Jemalong A17"
/db_xref="taxon:3880"
/clone_lib="MTE1"
/note="Vector: pIndigoBAC ; Site 1: EcoRI ; Site 2: EcoRI
; Debelle F. and Chalhou B.-Genoscope sequence ID :
mte1-35C7RM1"

ORIGIN
Query Match 100.0%; Score 16; DB 9; Length 567;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCAAGGTTTACCC 16
Db 155 ACCCCAAGGTTTACCC 170

RESULT 5
LOCUS CR348467 606 bp DNA linear GSS 04-MAR-2004
DEFINITION Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago
truncatula, genomic survey sequence.
ACCESSION CR348467
VERSION CR348467.1 GI:45120986
KEYWORDS GSS.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 606)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2004) Genoscope - Centre National de Sequencage :

```

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

## FEATURES

## source

1. .606  
/organism="Medicago truncatula"  
/mol\_type="genomic DNA"  
/cultivar="Jemalong A17"  
/db\_xref="taxon:3880"  
/clone\_lib="MTE1"  
/notes="Vector: pIndigoBAC ; Site 1: EcoRI ; Site 2: EcoRI  
; Debelle F. and Chalhou B.-Genoscope sequence ID :  
mte1-82N15FM1"

## ORIGIN

Query Match 100.0%; Score 16; DB 9; Length 606;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCCAGGTTTACCC 16

Db 172 ACCCCAGGTTTACCC 187

## RESULT 6

## CB986444

DEFINITION CB986444 759 bp mRNA linear EST 01-MAY-2003  
IMAGE:30329411 5', mRNA sequence.

## ACCESSION

## CB986444

## VERSION

## CB986444.1

## KEYWORDS

## EST.

## SOURCE

## Homo sapiens (human)

## ORGANISM

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

## NIH-MGC http://mgc.nci.nih.gov/.

## NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)

## UNPUBLISHED (1999)

## CONTACT: Robert Strausberg, Ph.D.

## Email: cgaabs-r@mail.nih.gov

## Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits

## CDNA Library Preparation: CLONTECH Laboratories, Inc.

## DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

## Clone distribution by: Agencourt Bioscience Corporation

## found through the I.M.A.G.E. Consortium/LLNL at:

## http://image.llnl.gov

## Plate: NDCM141 row: j column: 12

## High quality sequence stop: 468.

## Location/Qualifiers

## 1. .759

## /organism="Homo sapiens"

## /mol\_type="mRNA"

## /db\_xref="taxon:9606"

## /clone="IMAGE:30329411"

## /lab\_host="DH10B (T1 phage-resistant)"

## /clone\_lib="NIH\_MGC\_184"

## /notes="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site\_1:

## SfiI (ggcgattatggcc); Site\_2: SfiI (ggcgcttcggcc);

## Library is oligo-dT primed and directionally cloned. cDNA

## was prepared from a glandular pool of tissues from thyroid,

## parathyroid, adrenal, cortex and pineal gland. 5' and 3'

## adaptors were used in cloning as follows: 5' adaptor

## sequence: 5'-CAGGCCATATGCGC-3' and 3' adaptor sequence:

## 5'-ATTCTAGAGCCGAGCGCGCATG-dt(30)BN-3' (where B = A,

## C, G and N = A, C, G, or T). Average insert size 1.38

## kb (range 0.60-3.5 kb). 15/15 colonies contained inserts

## by PCR. This library was enriched for full-length clones

## and was constructed by Clontech Laboratories (Palo Alto,

## CA). Note: this is a NIH\_MGC Library."

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## 16; Conservative

## 0; Mismatches

## 0; Indels

## 0; Gaps

## 0;

## Qy

## 1

## ACCCCAAGGTTTACCC

## 16

## Db

## 675

## ACCCCAAGGTTTACCC

## 690

## RESULT 7

## AZ553628

## LOCUS

## DEFINITION

## RPCI-23-209L13-TV RPCI-23

## Mus musculus genomic clone

## genomic survey sequence.

## ACCESSION

## AZ553628

## VERSION

## AZ553628.1

## KEYWORDS

## GSS.

## SOURCE

## Mus musculus (house mouse)

## ORGANISM

## Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

## REFERENCE

## 1 (bases 1 to 788)

## AUTHORS

## Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,

## Akinret, B., Levine, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de

## Jong, P. and Fraser, C.M.

## Mouse BAC End Sequences from Library RPCI-23

## Unpublished (1999)

## CONTACT: Shaying Zhao

## Department of Eukaryotic Genomics

## The Institute for Genomic Research

## 9712 Medical Center Dr., Rockville, MD 20850, USA

## Tel: 301 838 0200

## Fax: 301 838 0208

## Email: szhao@tigr.org

## Clones are derived from the mouse BAC library RPCI-23. For BAC

## library availability, please contact Pieter de Jong

## (pieter@dejong.med.buffalo.edu). Clones may be purchased from

## BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)

## or from Resea ch Genetics (info@resgen.com). BAC end page:

## http://www.tigr.org/tdb/bac\_end/mouse/bac\_end\_intro.html

## Plate: 209 row: L column: 13

## Seq primer: T7

## Class: BAC ends.

## Location/Qualifiers

## 1. .788

## /organism="Mus musculus"

## /mol\_type="genomic DNA"

## /strain="C57BL/6J"

## /db\_xref="taxon:10090"

## /clone="RPCI-23-209L13"

## /sex="Female"

## /lab\_host="DH10B"

## /clone\_lib="RPCI-23"

## /notes="Organ: Kidney/Brain; Vector: pBACE3.6; Site\_1:

## EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or

## brain genomic DNA was isolated and partially digested

## with a combination of EcoRI and EcoRI Methyase. Size

## selected DNA was cloned into the pBACE3.6 vector at the

## EcoRI sites. The ligation products were transformed into

## DH10B electrocompetent cells (BRL Life Technologies)."

## ORIGIN

## Query Match

## Best Local Similarity

## Matches

## 16; Conservative

## 0; Mismatches

## 0; Indels

## 0; Gaps

## 0;

## Qy

## 1

## ACCCCAAGGTTTACCC

## 16

## Db

## 584

## ACCCCAAGGTTTACCC

## 599

## RESULT 8

## BU931691

```

LOCUS       BU931691                1017 bp    mRNA    linear    EST 18-OCT-2002
DEFINITION  AGENCOURT_10472385 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE:6671925 5', mRNA sequence.
ACCESSION   BU931691
VERSION     BU931691.1    GI:24120510
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1017)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2951 row: j column: 21
            High quality sequence stop: 579.

FEATURES             source
  Location/Qualifiers
    1..1017
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:6671925"
     /tissue_type="teratocarcinoma, cell line"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH MGC 109"
     /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
            XhoI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GCCACGAG(G). Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a
            NIH_MGC Library."

ORIGIN
Query Match      100.0%; Score 16; DB 5; Length 1017;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ACCCAAGGTTTACCC 16
        |||||||||||
        798  ACCCAAGGTTTACCC 813

Db

RESULT 9
CD888930/c
LOCUS       CD888930                391 bp    mRNA    linear    EST 14-JUL-2003
DEFINITION  G118.110J07F010720 G118 Triticum aestivum cDNA clone G118110707,
mRNA sequence.
ACCESSION   CD888930
VERSION     CD888930.1    GI:32656789
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poideae; Triticeae; Triticum.
            1 (bases 1 to 391)
REFERENCE   1
AUTHORS     Genoplante.
TITLE       Genoplante, a major partnership french program in plant genomics
JOURNAL     Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France

```

```

Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).

FEATURES             source
  Location/Qualifiers
    1..391
     /organism="Triticum aestivum"
     /mol_type="mRNA"
     /cultivar="recital"
     /db_xref="taxon:4565"
     /clone="G118110J07"
     /tissue_type="grain (118 degrees per day after
     pollination)"
     /clone_lib="G118"

ORIGIN
Query Match      93.8%; Score 15; DB 6; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  CCCCAAGGTTTACCC 16
        |||||||||||
        274  CCCCAAGGTTTACCC 260

Db

RESULT 10
CA706577/c
LOCUS       CA706577                412 bp    mRNA    linear    EST 26-NOV-2002
DEFINITION  wdk1c.pk024.j5 wdk1c Triticum aestivum cDNA clone wdk1c.pk024.j5 5',
end, mRNA sequence.
ACCESSION   CA706577
VERSION     CA706577.1    GI:25428370
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poideae; Triticeae; Triticum.
            1 (bases 1 to 412)
REFERENCE   1
AUTHORS     Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
            DuPont Wheat cDNA Sequence
            Unpublished (2002)
            Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.

FEATURES             source
  Location/Qualifiers
    1..412
     /organism="Triticum aestivum"
     /mol_type="mRNA"
     /db_xref="taxon:4565"
     /clone="wdk1c.pk024.j5"
     /tissue_type="kernel"
     /clone_lib="wdk1c"
     /note="Vector: pBluescript SK+; Wheat (Triticum aestivum
            L.) developing kernel, 3 days after anthesis."

ORIGIN
Query Match      93.8%; Score 15; DB 6; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  CCCCAAGGTTTACCC 16
        |||||||||||
        293  CCCCAAGGTTTACCC 279

```



## ORIGIN

Query Match 93.8%; Score 15; DB 6; Length 496;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCCAAGGTTTACCC 16  
 Db 60 CCCCAAGGTTTACCC 74

RESULT 14  
 BP641522/c  
 LOCUS BP641522.1 499 bp mRNA linear EST 27-JUN-2004  
 DEFINITION BP641522.1 Arabidopsis thaliana cDNA clone RAFL19-55-I18 3',  
 mRNA sequence.

ACCESSION BP641522  
 VERSION BP641522.1 GI:49292992  
 KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 499)

REFERENCE  
 AUTHORS Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,  
 Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M.,  
 Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y.,  
 Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.

TITLE Functional annotation of a full-length Arabidopsis cDNA collection

JOURNAL Science 296 (5565), 141-145 (2002)

MEDLINE 21932900

PUBMED 11910074

COMMENT Contact: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: mseki@tc.riken.go.jp

reversed clone; Please visit our web site

(<http://pfweb.gsc.riken.go.jp/>) for further details.

FEATURES

source

1..499  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL19-55-I18"  
 /tissue\_type="mixture of silique and flower"  
 /lab\_host="DH10B"  
 /clone\_lib="RAFL19"  
 /note="Site\_1: BamHI; Site\_2: SalI; Subtraction Library"

ORIGIN

Query Match 93.8%; Score 15; DB 5; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCCCAAGGTTTACCC 16  
 Db 359 CCCCAAGGTTTACCC 345

RESULT 15  
 BJ538102/c  
 LOCUS BJ538102.1 528 bp mRNA linear EST 09-AUG-2002  
 DEFINITION BJ538102.1 MF01SSB cDNA Oryzias latipes cDNA clone MF01SSB008H23 3',  
 mRNA sequence.

ACCESSION BJ538102  
 VERSION BJ538102.1 GI:22196914  
 KEYWORDS EST.

SOURCE Oryzias latipes (Japanese medaka)  
 ORGANISM Oryzias latipes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
 Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE  
 AUTHORS Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.  
 TITLE Medaka EST Project in Takeda's lab  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshini@genes.nig.ac.jp.

FEATURES  
 source

1..528  
 /organism="Oryzias latipes"  
 /mol\_type="mRNA"  
 /strain="Hd-rR"  
 /db\_xref="taxon:8090"  
 /clone="MF01SSB008H23"  
 /sex="mixture of female and male"  
 /tissue\_type="whole embryo"  
 /dev\_stage="segmentation stage 20 - 25"  
 /clone\_lib="MF01SSB cDNA"

ORIGIN

Query Match 93.8%; Score 15; DB 4; Length 528;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCCAAGGTTTACC 15  
 Db 525 ACCCCAAGGTTTACC 511

Search completed: May 16, 2005, 06:03:51  
 Job time : 1814 secs